

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted, and is derived by analysis of the total score distribution.

- esp. useful as blood plasma expanders

Disclosure: fig 2, 20pp. English

Nature Protein of human serum albumin (see corresp. AAN90C2a).  
Used to make new N-terminal fragments which are used as plasma  
expanders, or as substitutes for HSA or BSA in tissue culture  
media.

Updated on 25-MAR-2003 to correct PA field:

Sequence 585 AA:

Query Match 100.0% Score 3103. ID 10: Length 585.  
Ref: UniProt: P02768.1. ID 10: Length 585.  
Matches 585: Conservative C: Mismatches 0: Gaps 0:

```

1  DAHSEVAHREKLGSENFALVLAFAQYLOQCPPEHVKVNEVFETAKTVADSSAE 60
2  DAHSEVAHREKLGSENFALVLAFAQYLOQCPPEHVKVNEVFETAKTVADSSAE 60
61  NCKSLHTLFOKLCTVATLRETYGEMACCCAKGPERNECTJQKMDNPNLPRVPEV 120
121  DMVCTAFHNEETFLKYLYVEASRPHTVAPPELLFAKRYKAAFTCCOAKKAQLP 180
181  KLCLEDECKKASSAKORUKCASQKGFGEAFKAMAVARLSQRFPAEVSCLYDLTK 240
181  KLCLEDECKKASSAKORUKCASQKGFGEAFKAMAVARLSQRFPAEVSCLYDLTK 240
241  VHTECGGLLECCADRADAKYICENQDSISSKKECEKPLLEKSHCIAEVNDMPA 300
241  VHTECGGLLECCADRADAKYICENQDSISSKKECEKPLLEKSHCIAEVNDMPA 300
301  DLPFLAADFVSKNCKVYAEKADVFLGMFLYEVAHRHPVSVLLRLAKTYETLEK 360
301  DLPFLAADFVSKNCKVYAEKADVFLGMFLYEVAHRHPVSVLLRLAKTYETLEK 360
421  PTLVNSRNLKYGSKCNENKMYGAEYLSLVANGLVHENTPVSQATKCTTES 480
421  PTLVNSRNLKYGSKCNENKMYGAEYLSLVANGLVHENTPVSQATKCTTES 480
541  KSQLVAMDCFAAFVSKCCNADKMTCTFAESGXKLVAASQAALGL 585
541  KSQLVAMDCFAAFVSKCCNADKMTCTFAESGXKLVAASQAALGL 585

```

RESULT 2

AAN90C2a

AAN90C2a standard, protein: 585 AA.

AAN90C2a

AAN90C2a

28-OCT-1950 (first entry)

Human serum albumin

Human serum albumin: HSA-A, Yeast, ds.

Homo sapiens.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

NCBI: 17364-A.

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

28-OCT-1950

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

26-OCT-1958

[illegible]











X	Human; albumin; HA; fusion protein; therapeutic Protein; urinary;
X	immune system disorder; transplant rejection; blood related disorder;
X	myocardial infarction; hyperproliferative disorder; glomerulonephritis;
X	childhood acute myeloid leukemia; cardiovascular disorder; arrhythmia;
X	respiratory disorder; gene therapy; non-alcoholic thinitis; neoplastic;
X	endocrine disease; Alzheimer's disease; reproductive system disorder;
X	neurotic disorder; rheocystochrona; infectious disease; antiarrhythmia;
X	dysrhythmia; gastrointestinal disorder; irritable bowel syndrome; synchitis;
X	cardiac; antineoplastic; immunosuppressive; neuroprotective;
X	renal disorder.
X	
X	Homo sapiens.
S	
S	
I	Key Location/Qualifiers
T	Corain 54..61
L	/label= Loop_I
L	Domain /label= Loop_I
D	92..100
C	/label= Loop_II:
C	170..176
D	/label= Loop_IV
D	247..252
D	/label= Loop_V
D	266..277
D	/label= Loop_VI
D	362..368
D	/label= Loop_VIII
D	439..447
D	/label= Loop_IX
D	461..475
D	/label= Loop_X
D	475..494
D	/label= Loop_XI
D	505..514
D	/label= Loop_XII:
X	
X	WO200179443-A2
X	
X	25-OCT-2001.
X	
X	12-APR-2001; 2001MO-US:1924
X	
X	12-APR-2000; 2000US-22738P.
X	12-APR-2000; 2000US-23941P.
X	2-DEC-2000; 2000US-25691P
X	
X	(HUMA-) HUMAN GENOME SCI INC.
A	
K	Rosen CA, Haseltine WA;
I	
I	NF2; 2001-216754/71.
X	N-P50B; AAC21638
X	
X	Aluminum fusion proteins comprising a therapeutic protein and albumin,
X	useful in treating systemic disorders (e.g., transplamt,
X	rejection), blood-related disorders (e.g., myocardial infarction) and
X	hyperproliferative disorders.
X	
X	Claim 1, Fig 9; 38pp; Eng:16t.
X	
X	The invention relates to albumin fusion proteins comprising therapeutic
X	protein and human albumin (HA). Therapeutic protein fused to albumin
X	have an extended half-life. The albumin fusion proteins are useful in
X	treatng various diseases and disorders such as immune system disorders,
X	blood related disorders (e.g., myocardial infarction), hyperproliferative
X	disorders (e.g., childhood acute myeloid leukaemia), renal disorders
X	(e.g., glomerulonephritis), cardiovascular disorders (e.g., arrhythmias),

[illegible]









MOLECULE TYPE: protein  
HYDROPHOBICITY: 10  
ORGANISM: Homo sapiens  
FEATURES:  
NAME/KEY: Region  
LOCATION: 349-419  
OTHER INFORMATION: /note="Alternative C-termini of  
OTHER INFORMATION: HSA(1-6)"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 545  
OTHER INFORMATION: /note="amino acid sequence of  
OTHER INFORMATION: natural HSA"  
US-08-153-759-14  
Query Match 130.04; Score 3103; DB 1; Length 585;  
Best Local Similarity 130.04; Pred. No. 9, 30-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 1 DAHSEVAHFKDSEKFAVLAFAGYVQCQPFEDHVKLVNEVTEFAKTCVADESAR 60  
DB 1 DAHSEVAHFKDSEKFAVLAFAGYVQCQPFEDHVKLVNEVTEFAKTCVADESAR 60  
CY 6: NDKSLHTLFCGLCTVATLETETGEMADCCAKQEPENECVQNTDNPPLTPVWBPV 120  
DB 6: NDKSLHTLFCGLCTVATLETETGEMADCCAKQEPENECVQNTDNPPLTPVWBPV 120  
CY 6: NDKSLHTLFCGLCTVATLETETGEMADCCAKQEPENECVQNTDNPPLTPVWBPV 120  
DB 6: NDKSLHTLFCGLCTVATLETETGEMADCCAKQEPENECVQNTDNPPLTPVWBPV 120  
CY 121 DVMCTAFHNEETFLKXYVEARRHPYFAPELLFFAKYKAAPTECCCAAKKACLP 180  
DB 121 DVMCTAFHNEETFLKXYVEARRHPYFAPELLFFAKYKAAPTECCCAAKKACLP 180  
CY 121 DVMCTAFHNEETFLKXYVEARRHPYFAPELLFFAKYKAAPTECCCAAKKACLP 180  
DB 121 DVMCTAFHNEETFLKXYVEARRHPYFAPELLFFAKYKAAPTECCCAAKKACLP 180  
CY 13: KDELDEKASAKQKQKASQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 240  
DB 13: KDELDEKASAKQKQKASQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 240  
CY 241 VHTTECHGELLECAADRALAKYICENQDSISSKKECKEFLKSHHCIAEVNDMPA 300  
DB 241 VHTTECHGELLECAADRALAKYICENQDSISSKKECKEFLKSHHCIAEVNDMPA 300  
CY 301 DLSLAADPFESKQVACVYARADVFLQVLYVYARHPDYVYLLRLAKTYETTLK 360  
DB 301 DLSLAADPFESKQVACVYARADVFLQVLYVYARHPDYVYLLRLAKTYETTLK 360  
CY 361 GAAAPHECAKVFDEKFEVLEVEONLKVONCELFQQLGEVYKFNALVYTKVPOVST 420  
DB 361 GAAAPHECAKVFDEKFEVLEVEONLKVONCELFQQLGEVYKFNALVYTKVPOVST 420  
CY 421 PTUVEVRNLKVSCKQ 480  
DB 421 PTUVEVRNLKVSCKQ 480  
CY 481 DVMCTAFHNEETFLKXYVEARRHPYFAPELLFFAKYKAAPTECCCAAKKACLP 540  
DB 481 DVMCTAFHNEETFLKXYVEARRHPYFAPELLFFAKYKAAPTECCCAAKKACLP 540  
CY 541 KEGKAMQDFPAKVEKQ 596  
DB 541 KEGKAMQDFPAKVEKQ 596

RESULT 2  
US-08-732-572-2  
Sequence 2, Application US/08/732-2  
Patent No. 556386  
GENERAL INFORMATION  
APPLICANT: Genentech, Inc.  
TITLE OF INVENTION: Variant Stimulation and Modified Amino Acids  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS  
ADDRESSEE: Genentech, Inc.  
STREET: 1020 First Avenue

CITY: Kind of House-A  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19106-1112  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WinGsoft Word 6.0  
CURRENT APPLICATION DATA  
APPLICATION NUMBER: US/08/701,572  
FILING DATE: 1-NOV-1996  
CLASSIFICATION: C12N 1/21  
PRIORITY INFORMATION: 0  
PRIORITY NUMBER: AC 95/23857  
FILING DATE: 1-MAR-1994  
APPLICATION NUMBER: GB 9404270.2  
FILING DATE: 5-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Naomi Blasius  
REGISTRATION NUMBER: 34,164  
REFERENCE/DOCKET NUMBER: CE0114 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610/4984224  
TELEFAX: 610/4984224  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS  
LENGTH: 585 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-732-572-2  
Query Match 130.04; Score 3103; DB 2; Length 585;  
Best Local Similarity 130.04; Pred. No. 9, 30-287;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 1 DAHSEVAHFKDSEKFAVLAFAGYVQCQPFEDHVKLVNEVTEFAKTCVADESAR 60  
DB 1 DAHSEVAHFKDSEKFAVLAFAGYVQCQPFEDHVKLVNEVTEFAKTCVADESAR 60  
CY 6: NDKSLHTLFCGLCTVATLETETGEMADCCAKQEPENECVQNTDNPPLTPVWBPV 120  
DB 6: NDKSLHTLFCGLCTVATLETETGEMADCCAKQEPENECVQNTDNPPLTPVWBPV 120  
CY 6: NDKSLHTLFCGLCTVATLETETGEMADCCAKQEPENECVQNTDNPPLTPVWBPV 120  
DB 6: NDKSLHTLFCGLCTVATLETETGEMADCCAKQEPENECVQNTDNPPLTPVWBPV 120  
CY 121 DVMCTAFHNEETFLKXYVEARRHPYFAPELLFFAKYKAAPTECCCAAKKACLP 180  
DB 121 DVMCTAFHNEETFLKXYVEARRHPYFAPELLFFAKYKAAPTECCCAAKKACLP 180  
CY 121 DVMCTAFHNEETFLKXYVEARRHPYFAPELLFFAKYKAAPTECCCAAKKACLP 180  
DB 121 DVMCTAFHNEETFLKXYVEARRHPYFAPELLFFAKYKAAPTECCCAAKKACLP 180  
CY 13: KDELDEKASAKQKQKASQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 240  
DB 13: KDELDEKASAKQKQKASQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 240  
CY 241 VHTTECHGELLECAADRALAKYICENQDSISSKKECKEFLKSHHCIAEVNDMPA 300  
DB 241 VHTTECHGELLECAADRALAKYICENQDSISSKKECKEFLKSHHCIAEVNDMPA 300  
CY 301 DLSLAADPFESKQVACVYARADVFLQVLYVYARHPDYVYLLRLAKTYETTLK 360  
DB 301 DLSLAADPFESKQVACVYARADVFLQVLYVYARHPDYVYLLRLAKTYETTLK 360  
CY 361 GAAAPHECAKVFDEKFEVLEVEONLKVONCELFQQLGEVYKFNALVYTKVPOVST 420  
DB 361 GAAAPHECAKVFDEKFEVLEVEONLKVONCELFQQLGEVYKFNALVYTKVPOVST 420  
CY 421 PTUVEVRNLKVSCKQ 480  
DB 421 PTUVEVRNLKVSCKQ 480  
CY 481 DVMCTAFHNEETFLKXYVEARRHPYFAPELLFFAKYKAAPTECCCAAKKACLP 540  
DB 481 DVMCTAFHNEETFLKXYVEARRHPYFAPELLFFAKYKAAPTECCCAAKKACLP 540  
CY 541 KEGKAMQDFPAKVEKQ 596  
DB 541 KEGKAMQDFPAKVEKQ 596

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-433-037-4

Query Match: 99.98; Score 3099; DB 1; Length 609;
Best Local Similarity 99.88; Pred. No. 2.4e-286;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPAOYLQCCPREDHVKLNVETEFAKTCVADESAE 60
DB 25 DAHSEVAHRFKDGLGEENFKALVLIAPAOYLQCCPREDHVKLNVETEFAKTCVADESAE 94
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRVLRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRVLRPEV 144
QY 121 DVMTCTAHDNEETFLKKYLVEIARRHPYFYAPPELLFPAKYKAAFTCCCOADKAAACLP 180
DB 145 DVMTCTAHDNEETFLKKYLVEIARRHPYFYAPPELLFPAKYKAAFTCCCOADKAAACLP 204
QY 181 KLDELREGKASSAKORLKCASLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240
DB 205 KLDELREGKASSAKORLKCASLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYBYARRHPDYSVVLRLRAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYBYARRHPDYSVVLRLRAKTYETTLK 384
QY 361 CAADAPHECYAKVFDEFKPLVEEPQNLIKQNCBLFEQLGYKFNALLVRYTKKVPQVST 420
DB 385 CAADAPHECYAKVFDEFKPLVEEPQNLIKQNCBLFEQLGYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
DB 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 504
QY 481 LVNRRPFCFSALEVDVETVVPKEFNAETTFPHADCTLSEKERQIKKQTALVELVKHKPKAT 540
DB 505 LVNRRPFCFSALEVDVETVVPKEFNAETTFPHADCTLSEKERQIKKQTALVELVKHKPKAT 564
QY 541 KEQLKAVMDQFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDQFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609

RESULT 10
US-08-897-956A-2
; Sequence 2, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CPA
; CURRENT APPLICATION NUMBER: US/08/897.956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022.689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PPT
; ORGANISM: Homo Sapiens
US-08-897-956A-2
```

```
Query Match: 99.98; Score 3099; DB 4; Length 609;
Best Local Similarity 99.88; Pred. No. 2.4e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPAOYLQCCPREDHVKLNVETEFAKTCVADESAE 60
DB 25 DAHSEVAHRFKDGLGEENFKALVLIAPAOYLQCCPREDHVKLNVETEFAKTCVADESAE 94
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRVLRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRVLRPEV 144
QY 121 DVMTCTAHDNEETFLKKYLVEIARRHPYFYAPPELLFPAKYKAAFTCCCOADKAAACLP 180
DB 145 DVMTCTAHDNEETFLKKYLVEIARRHPYFYAPPELLFPAKYKAAFTCCCOADKAAACLP 204
QY 181 KLDELREGKASSAKORLKCASLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240
DB 205 KLDELREGKASSAKORLKCASLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYBYARRHPDYSVVLRLRAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYBYARRHPDYSVVLRLRAKTYETTLK 384
QY 361 CAADAPHECYAKVFDEFKPLVEEPQNLIKQNCBLFEQLGYKFNALLVRYTKKVPQVST 420
DB 385 CAADAPHECYAKVFDEFKPLVEEPQNLIKQNCBLFEQLGYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 480
DB 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTES 504
QY 481 LVNRRPFCFSALEVDVETVVPKEFNAETTFPHADCTLSEKERQIKKQTALVELVKHKPKAT 540
DB 505 LVNRRPFCFSALEVDVETVVPKEFNAETTFPHADCTLSEKERQIKKQTALVELVKHKPKAT 564
QY 541 KEQLKAVMDQFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDQFAAFVEKCKCKADDKETCFABEGKKLVAASQAALGL 609

RESULT 11
PCT-US95-04375-3
; Sequence 3, Application PC/TUS9504075
; GENERAL INFORMATION:
; APPLICANT: AMGEN INC.
; TITLE OF INVENTION: Afamir: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESS: Amgen Center, Patent Operations/RR
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
```

```
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
PCT-US95-04075-3

Query Match      99.9%; Score 3099; DB 5; Length 609;
Best Local Similarity 99.8%; Pred. No. 2.4e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAKSEVAHRRFKDLGEENFKALVLIAPAFYQYQQCPFEHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAKSEVAHRRFKDLGEENFKALVLIAPAFYQYQQCPFEHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNPKNLRLVRREV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNPKNLRLVRREV 144
QY 121 DVMCTAFHNDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180
DB 145 DVMCTAFHNDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 204
QY 181 KLDELRDGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEPAEVSKLVTDLTK 240
DB 205 KLDELRDGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEPAEVSKLVTDLTK 264
QY 241 VHTCCHGDLLECADRADLAKYICENODSISSKLKECEKPLLEKSHCAEVENDEMPA 300
DB 265 VHTCCHGDLLECADRADLAKYICENODSISSKLKECEKPLLEKSHCAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGNFLYEVARRHPDYSVLLRLAKTYETTLKX 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFLGNFLYEVARRHPDYSVLLRLAKTYETTLKX 384
QY 361 CAAADPHECYAKVDFEPKPLVEEPQNLKQNCLEPQGEYKFNQALLVRYTKVPQVST 420
DB 385 CAAADPHECYAKVDFEPKPLVEEPQNLKQNCLEPQGEYKFNQALLVRYTKVPQVST 444
QY 421 PTLVEVSRNLGVGSKCKGHPKAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGVGSKCKGHPKAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALAEVDYTPVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 540
DB 505 LVNRRPCFSALAEVDYTPVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 564
QY 541 KEQLKAVMDDFAAVFVEKCKCKADDKETCTFAEBGKGLVAASQAALGL 585
DB 565 KEQLKAVMDDFAAVFVEKCKCKADDKETCTFAEBGKGLVAASQAALGL 609
```

```
RESULT 12
US-08-897-956A-3
; Sequence 3, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CPA
; CURRENT APPLICATION NUMBER: US/08/897,956A
; PRIOR FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 978
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion polypeptide
```

```
US-08-897-956A-3
Query Match      99.7%; Score 3095; DB 4; Length 978;
Best Local Similarity 99.6%; Pred. No. 1.1e-285;
Matches 583; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAKSEVAHRRFKDLGEENFKALVLIAPAFYQYQQCPFEHVKLVNEVTEFAKTCVADESAAE 60
DB 212 DAKSEVAHRRFKDLGEENFKALVLIAPAFYQYQQCPFEHVKLVNEVTEFAKTCVADESAAE 271
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNPKNLRLVRREV 120
DB 272 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPENECFLQHKDNPKNLRLVRREV 331
QY 121 DVMCTAFHNDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180
DB 332 DVMCTAFHNDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 391
QY 191 KLDELRDGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEPAEVSKLVTDLTK 240
DB 392 KLDELRDGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEPAEVSKLVTDLTK 451
QY 241 VHTCCHGDLLECADRADLAKYICENODSISSKLKECEKPLLEKSHCAEVENDEMPA 300
DB 452 VHTCCHGDLLECADRADLAKYICENODSISSKLKECEKPLLEKSHCAEVENDEMPA 511
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGNFLYEVARRHPDYSVLLRLAKTYETTLKX 360
DB 512 DLPSLAADFVESKDVCKNYAEAKDVFLGNFLYEVARRHPDYSVLLRLAKTYETTLKX 571
QY 361 CAAADPHECYAKVDFEPKPLVEEPQNLKQNCLEPQGEYKFNQALLVRYTKVPQVST 420
DB 572 CAAADPHECYAKVDFEPKPLVEEPQNLKQNCLEPQGEYKFNQALLVRYTKVPQVST 631
QY 421 PTLVEVSRNLGVGSKCKGHPKAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 632 PTLVEVSRNLGVGSKCKGHPKAKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 691
QY 481 LVNRRPCFSALAEVDYTPVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 540
DB 692 LVNRRPCFSALAEVDYTPVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 751
QY 541 KEQLKAVMDDFAAVFVEKCKCKADDKETCTFAEBGKGLVAASQAALGL 584
DB 752 KEQLKAVMDDFAAVFVEKCKCKADDKETCTFAEBGKGLVAASQAALGL 795
```

RESULT 13

US-08-448-196A-3

; Sequence 3, Application US/08448196A

; Patent No. 5780594

; GENERAL INFORMATION:

; APPLICANT: CARTER, DANIEL C.

; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS

; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR

; TITLE OF INVENTION: RELATED PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NASA

; STREET: MARSHALL SPACE FLIGHT CENTER

; CITY: HUNTSVILLE

; STATE: ALABAMA

; COUNTRY: USA

; ZIP: 35812

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/448,196A

; FILING DATE: 23-MAY-1995

; CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:  
 NAME: BROAD JR., ROBERT L.  
 REGISTRATION NUMBER: 18,757  
 REFERENCE/DOCKET NUMBER: XX/WFS-28422-2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 205-544-0021  
 TELEFAX: 205-544-0258  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 585 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEetical: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 US-08-448-196A-3

```

Query Match          99.7%   Score 3093;   DB 1;   Length 585;
Best Local Similarity 99.7%   Pred. No. 8.3e-286;
Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 1 DAHSEVAHREFKDLGGEENFKALVLIIFAQYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHREFKDLGGEENFKALVLIIFAQYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60

Cy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

Cy 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLPL 180
Db 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLPL 180

Cy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQFPKAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQFPKAEFAEVSKLVTDLTK 240

Cy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDENPA 300
Db 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDENPA 300

Cy 301 DPLSLAADFVESKDVCKNYAEAKDVFGLGMFLVEYARRHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DPLSLAADFVESKDVCKNYAEAKDVFGLGMFLVEYARRHPDYSVLLRLAKTYETTTLEKC 360

Cy 361 CAADHPHECYAKVDFDEPKLVBEPPQLIKONCELFQKLGVEYKFNALLVRYTKVPQVST 420
Db 361 CAADHPHECYAKVDFDEPKLVBEPPQLIKONCELFQKLGVEYKFNALLVRYTKVPQVST 420

Cy 421 PTLVEVRNLGKVGSKCKHPKAEKMPCAEDYLSVWNLQCLVLEKTPVSDRVTKCCCTES 480
Db 421 PTLVEVRNLGKVGSKCKHPKAEKMPCAEDYLSVWNLQCLVLEKTPVSDRVTKCCCTES 480

Cy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540

Cy 541 KEQLKAVMDDFAAFEVKCKADDKETCFABEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFEVKCKADDKETCFABEGKKLVAASQAALGL 585

```

## RESULT 14

US-08-984-176-1  
 Sequence 1, Application US/08984176  
 Patent No. 5948609  
 GENERAL INFORMATION:  
 APPLICANT: CARTER, DANIEL C  
 APPLICANT: HO, JOSEPH X  
 APPLICANT: RUKER, FLORIAN  
 TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT  
 TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER

FILE REFERENCE: 08/984,176  
 CURRENT APPLICATION NUMBER: US/08/984,176  
 CURRENT FILING DATE: 1997-12-03  
 NUMBER OF SEQ ID NOS: 1  
 SOFTWARE: Patentit Ver. 2.0  
 SEQ ID NO 1  
 LENGTH: 585  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-08-984-176-1

```

Query Match          99.7%   Score 3093;   DB 2;   Length 585;
Best Local Similarity 99.7%   Pred. No. 8.3e-286;
Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 1 DAHSEVAHREFKDLGGEENFKALVLIIFAQYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHREFKDLGGEENFKALVLIIFAQYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60

Cy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

Cy 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLPL 180
Db 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLPL 180

Cy 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQFPKAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQFPKAEFAEVSKLVTDLTK 240

Cy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDENPA 300
Db 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDENPA 300

Cy 301 DPLSLAADFVESKDVCKNYAEAKDVFGLGMFLVEYARRHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DPLSLAADFVESKDVCKNYAEAKDVFGLGMFLVEYARRHPDYSVLLRLAKTYETTTLEKC 360

Cy 361 CAADHPHECYAKVDFDEPKLVBEPPQLIKONCELFQKLGVEYKFNALLVRYTKVPQVST 420
Db 361 CAADHPHECYAKVDFDEPKLVBEPPQLIKONCELFQKLGVEYKFNALLVRYTKVPQVST 420

Cy 421 PTLVEVRNLGKVGSKCKHPKAEKMPCAEDYLSVWNLQCLVLEKTPVSDRVTKCCCTES 480
Db 421 PTLVEVRNLGKVGSKCKHPKAEKMPCAEDYLSVWNLQCLVLEKTPVSDRVTKCCCTES 480

Cy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540

Cy 541 KEQLKAVMDDFAAFEVKCKADDKETCFABEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFEVKCKADDKETCFABEGKKLVAASQAALGL 585

```

## RESULT 15

US-08-448-196A-5  
 Sequence 5, Application US/08448196A  
 Patent No. 5780594  
 GENERAL INFORMATION:  
 APPLICANT: CARTER, DANIEL C  
 TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS  
 TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR  
 TITLE OF INVENTION: RELATED PROTEINS  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: NASA  
 STREET: MARSHALL SPACE FLIGHT CENTER  
 CITY: HUNTSVILLE  
 STATE: ALABAMA  
 COUNTRY: USA  
 ZIP: 35812

COMPUTER READABLE FORM:

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent'n Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,196A
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: BROAD JR., ROBERT L.
REGISTRATION NUMBER: 18,757
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 205-544-0021
TELEFAX: 205-544-0258
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-448-196A-5

Query Match 79.2% Score 2458.5; DB 1; Length 583;
Best Local Similarity 75.8% Pred. No. 2.1e-225;
Matches 442; Conservative 70; Mismatches 70; Indels 1; Gaps 1

QY 1 DAHSEVAHRKQJGEBNFKAULVIAFAQYLQCPPEFDHVKLVNVEYFAKTCVADESA 60
DB 1 DTHKSEIAHRENDAGEKHFKGLVIVAFSoyLQCPPEFDHVKLVNVEYFAKTCVADESA 60
QY 61 NCDKSLHTLFGKLCVTATLETETGEMADCCAKOEPEENECFLQHKDDPNLPLEVGRPEV 120
DB 61 NCDKSLHTLFGKLCVTATLATYTGELADCCKEGEPERNECFILTKDDHPNLPKL-KPEF 119
QY 121 DVMCTAFHDNEETELKKYLYEIAHRHPYFYAPPELLFFAKYKAAFTCCQAAADKACLLP 180
DB 120 QAQCAAFCEQDFKGLKYLYEVARRHPYFYGPPELLFHAEBYKACFTCECCPADKACLIP 179
QY 181 KLDELRDGKASSAKORLCKASLOKQFERAFKAWAVARLSQRTPKAEPAEYVKLVTDLT 240
DB 180 KLDAUKERILLUSSAKERUKCSSFQNGFERAVKAWSVARLSQKTPKADPAEYVKLVTDLT 239
QY 241 VHTCECHGDLLECADRDADLAKYICENQDS:SSKLKBCCKPKLLEKSHCIAEYKNDMPA 300
DB 240 VHKECCHGDLLECADRDADLAKYICEHDS:SGKLKACCKPKLLOKSHCIAEYKEDDLP 299
QY 301 DLPSPAADPVSCKYCKNYAKAKOVLMGF:YBYARHPDYSVVLLRLAKYVETLEK 360
DB 300 DIPALAADPAEDKEIKGYIKDAKOVFGTFLYYSRRHPDYSVVSLLR-AKTYEATLEK 359
QY 361 CAAAPDHECYAKVDFEFPVPEPQNT:KQNCFLFEQLGEYKFNALLVRYTKKVPQST 420
DB 360 CAEADPPACYRTVPDF:PLVVEPKSLVKKNCDLFEVGEYDFQNALIVRYTKKAPQST 419
QY 421 PTLVEVSRNLGKVGSKCKHPFAKMPCAEDYLSVLNQICVLEKTPVSDRVTKCTES 480
DB 420 PTLVETRIGTLGKVGSRCKLPESRPLPCSENLHALA:NRLCVLEHKTTPVSEKITKCTDS 479
QY 481 LVNRRPCPSALEVDVETYPYKPFNAETETFHADI:CTLSEKERQIKKOTALVELVKKPKAT 540
DB 480 LAERRPCPSALELDGYPYKPFNAETETFHADI:CTLPEDEQ:KKQSALAEVKKPKAT 539
QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEBGKKLVAASGAAL 583
DB 540 KEQLKATVLGNFAFVAKCCGDEKCAEFAEBGPKLVAASGLAL 582

```

Search completed: October 27, 2003, 15:22:31

GenCore version 5.2.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

DM protein - protein search, using sw model

Run on: October 27, 2003, 15:21:13 ; Search time 72 Seconds  
(without alignments)  
1360.618 Million cell updates/sec

Title: US-09-833-117-18

Perfect score: 3103

Sequence: 1 DAHSEVAHRKDI/GEENFK.....TCFAEECKKLVAASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US05\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_FJBOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US0A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEX\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	585	10	US-09-929-552-2
2	3103	100.0	585	11	US-09-932-63-445
3	3103	100.0	585	11	US-09-984-010-26
4	3103	100.0	585	11	US-09-833-041-18
5	3103	100.0	585	12	US-10-153-604A-5
6	3103	100.0	585	12	US-09-833-117-18
7	3103	100.0	585	12	US-10-319-263-1
8	3103	100.0	585	12	US-10-319-263-2
9	3103	100.0	585	12	US-10-414-469-1
10	3103	100.0	585	12	US-10-414-469-2
11	3103	100.0	585	12	US-09-932-322-445
12	3103	100.0	585	12	US-10-413-831-1
13	3103	100.0	585	12	US-10-413-831-2
14	3103	100.0	585	14	US-10-153-064-5
15	3103	100.0	634	11	US-09-984-010-7

16	3103	100.0	609	11	US-09-939-039-370	Sequence 370, Appl
17	3103	100.0	609	12	US-10-153-604A-7	Sequence 7, Appl1
18	3103	100.0	609	12	US-10-365-623-23	Sequence 23, Appl1
19	3103	100.0	609	14	US-10-153-064-7	Sequence 7, Appl1
20	3103	100.0	610	10	US-09-984-186-2	Sequence 2, Appl1
21	3103	100.0	610	15	US-10-237-667-2	Sequence 2, Appl1
22	3103	100.0	610	15	US-10-237-708-2	Sequence 2, Appl1
23	3103	100.0	610	15	US-10-237-866-2	Sequence 2, Appl1
24	3103	100.0	610	15	US-10-237-871-2	Sequence 2, Appl1
25	3103	100.0	610	15	US-10-237-624-2	Sequence 2, Appl1
26	3103	100.0	651	12	US-10-153-604A-133	Sequence 133, App
27	3103	100.0	651	14	US-10-153-064-133	Sequence 133, App
28	3103	100.0	652	12	US-10-153-604A-132	Sequence 132, App
29	3103	100.0	652	14	US-10-153-064-132	Sequence 132, App
30	3103	100.0	653	12	US-10-153-604A-131	Sequence 131, App
31	3103	100.0	653	14	US-10-153-064-131	Sequence 131, App
32	3103	100.0	656	12	US-10-153-604A-130	Sequence 130, App
33	3103	100.0	656	14	US-10-153-064-130	Sequence 130, App
34	3103	100.0	676	12	US-10-153-604A-129	Sequence 129, App
35	3103	100.0	676	12	US-10-153-604A-127	Sequence 127, App
36	3103	100.0	676	14	US-10-153-064-127	Sequence 127, App
37	3103	100.0	677	12	US-10-153-604A-125	Sequence 125, App
38	3103	100.0	677	14	US-10-153-064-125	Sequence 125, App
39	3103	100.0	680	12	US-10-153-604A-123	Sequence 123, App
40	3103	100.0	680	14	US-10-153-064-123	Sequence 123, App
41	3103	100.0	787	10	US-09-984-186-16	Sequence 16, Appl
42	3103	100.0	787	15	US-10-237-667-16	Sequence 16, Appl
43	3103	100.0	787	15	US-10-237-708-16	Sequence 16, Appl
44	3103	100.0	787	15	US-10-237-866-16	Sequence 16, Appl
45	3103	100.0	787	15	US-10-237-866-16	Sequence 16, Appl

#### ALIGNMENTS

#### RESULT :

US-09-929-552-2  
; Sequence 2, Application US/0929552  
; Patent No. US20020123080A1  
; GENERAL INFORMATION:  
; APPLICANT: Sonnenschein, Carlos  
; Soto, Ana M.  
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Medien & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DCS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/929,552  
; FILING DATE: 14-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/769,746  
; FILING DATE: 19-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: MBRI-02584  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 585 amino acids

RESULT 2  
 S-09-932-613-445  
 Sequence 445, Application US/09932613  
 Publication No. US20030091565A1  
 GENERAL INFORMATION:  
 APPLICANT: Human Genome Sciences, Inc.  
 APPLICANT: Beitzler, James P.  
 APPLICANT: Potter, M. Daniel  
 APPLICANT: Fleming, Tony J.  
 APPLICANT: Rosen, Craig A.  
 TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON  
 FILE REFERENCE: DXX-025.1 PCT: DXX-025.1 US  
 CURRENT APPLICATION NUMBER: US/09/932,613  
 CURRENT FILING DATE: 2001-08-17  
 NUMBER OF SEQ ID NOS: 458  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 445  
 LENGTH: 585  
 TYPE: PRT  
 ORGANISM: HomoSapiens  
 S-09-932-613-445

; FILING DATE: 19-DEC-1996  
 ; INFORMATION FOR SEQ ID NO: 26:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 585 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: <Unknown>  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHEetical: NO  
 ; ANTI-SENSE: NO  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
 US-09-984-010-26

Query Match 100.0%; Score 3103; DB 11; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-269;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DAHSEVAHAFKDLGSENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
 DB 1 DAHSEVAHAFKDLGSENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFQHKDDNPNLRLVRPEV 120  
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFQHKDDNPNLRLVRPEV 120  
 QY 121 DMCTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKYKAAFTCCQAADKAACLLP 180  
 DB 121 DMCTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKYKAAFTCCQAADKAACLLP 180  
 QY 181 KLDELREGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRFPAEFAEVSKLVTDLTK 240  
 DB 181 KLDELREGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRFPAEFAEVSKLVTDLTK 240  
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300  
 DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVLLRLAKTYETTLK 360  
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVLLRLAKTYETTLK 360  
 QY 361 CAADPHCYAKVDFEFPKPLVEEPQNLIKNCELFEQLGEYKFNALLVRYTKKVPQVST 420  
 DB 361 CAADPHCYAKVDFEFPKPLVEEPQNLIKNCELFEQLGEYKFNALLVRYTKKVPQVST 420  
 QY 421 PTLVEVSRLKGVSKCKKHPKAPKMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
 DB 421 PTLVEVSRLKGVSKCKKHPKAPKMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
 QY 481 LVNRRPCFSALEVDVETVYVPKEFNAETTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540  
 DB 481 LVNRRPCFSALEVDVETVYVPKEFNAETTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540  
 QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGJ 585  
 DB 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGJ 585

RESULT 4  
 US-09-833-041-18  
 ; Sequence 18, Application US/09833041  
 ; Publication No. US20030125247A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen, Craig A.  
 ; TITLE OF INVENTION: Albumin Fusion Proteins  
 ; FILE REFERENCE: PF545  
 ; CURRENT APPLICATION NUMBER: US/09/833,041  
 ; CURRENT FILING DATE: 2001-04-12  
 ; PRIOR APPLICATION NUMBER: 60/229,358  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: 60/256,931  
 ; PRIOR FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/199,384  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 79  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 18  
 ; LENGTH: 585  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapiens  
 US-09-833-041-18  
 Query Match 100.0%; Score 3103; DB 11; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-269;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DAHSEVAHAFKDLGSENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
 DB 1 DAHSEVAHAFKDLGSENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFQHKDDNPNLRLVRPEV 120  
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFQHKDDNPNLRLVRPEV 120  
 QY 121 DMCTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKYKAAFTCCQAADKAACLLP 180  
 DB 121 DMCTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKYKAAFTCCQAADKAACLLP 180  
 QY 181 KLDELREGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRFPAEFAEVSKLVTDLTK 240  
 DB 181 KLDELREGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRFPAEFAEVSKLVTDLTK 240  
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300  
 DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVLLRLAKTYETTLK 360  
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVLLRLAKTYETTLK 360  
 QY 361 CAADPHCYAKVDFEFPKPLVEEPQNLIKNCELFEQLGEYKFNALLVRYTKKVPQVST 420  
 DB 361 CAADPHCYAKVDFEFPKPLVEEPQNLIKNCELFEQLGEYKFNALLVRYTKKVPQVST 420  
 QY 421 PTLVEVSRLKGVSKCKKHPKAPKMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
 DB 421 PTLVEVSRLKGVSKCKKHPKAPKMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480  
 QY 481 LVNRRPCFSALEVDVETVYVPKEFNAETTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540  
 DB 481 LVNRRPCFSALEVDVETVYVPKEFNAETTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540  
 QY 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGJ 585  
 DB 541 KEQLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGJ 585

RESULT 5  
 US-10-153-604A-5  
 ; Sequence 5, Application US/10:53604A  
 ; Publication No. US20030:43191A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bell et al.  
 ; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
 ; FILE REFERENCE: PF556  
 ; CURRENT APPLICATION NUMBER: US/10/153,604A  
 ; CURRENT FILING DATE: 2002-05-24  
 ; PRIOR APPLICATION NUMBER: 60/293,212  
 ; PRIOR FILING DATE: 2001-05-25  
 ; NUMBER OF SEQ ID NOS: 137  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 585  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapiens

US-10-153-604A-5

Query Match 100.0%; Score 3103; DB 12; Length 585;  
Best Local Similarity 100.0%; Pred. No. 2.4e-269;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRFKDLGEENFKALVLIAPAFQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRRFKDLGEENFKALVLIAPAFQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDPNLRLVRREV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDPNLRLVRREV 120

QY 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180  
DB 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180

QY 181 KLDELDEGKASSAKORLKASLQKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKORLKASLQKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240

QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAAKOVFLGMFLYEVARRHPDYVSVLLRLAKTYETTLEK 360  
DB 301 DLPSLAADFVESKDVCKNYAAKOVFLGMFLYEVARRHPDYVSVLLRLAKTYETTLEK 360

QY 361 CAADPHCEYAKVDFEPLVEEPONLIKONCELFEOIGEYKFNALLVRYTKVPQVST 420  
DB 361 CAADPHCEYAKVDFEPLVEEPONLIKONCELFEOIGEYKFNALLVRYTKVPQVST 420

QY 421 PTLVEVSRNLGVGSKCKHPEAKMPCAEYDLSVNLQCLVLEHKTVPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGVGSKCKHPEAKMPCAEYDLSVNLQCLVLEHKTVPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDYTPVKFNAETFTFHADICTLSEKEROIKKQALVELVVKHKPKAT 540  
DB 481 LVNRRPCFSALEVDYTPVKFNAETFTFHADICTLSEKEROIKKQALVELVVKHKPKAT 540

QY 541 KEQLKAVMDFFAAFEVKCKKADDETCFAEEGKKLVAAASQAALGL 585  
DB 541 KEQLKAVMDFFAAFEVKCKKADDETCFAEEGKKLVAAASQAALGL 585

## RESULT 6

US-09-833-117-18  
; Sequence 18, Application US/098331:7  
; Publication No. US20030171267A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Sadeghi, Homa  
; APPLICANT: Prior, Christopher P.  
; APPLICANT: Turner, Andrew J.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF543  
; CURRENT APPLICATION NUMBER: US/09/833,117  
; CURRENT FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229,358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256,931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199,384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 585  
; TYPE: PR  
; ORGANISM: Homo Sapiens

US-09-833-117-18

Query Match 100.0%; Score 3103; DB 12; Length 585;  
Best Local Similarity 100.0%; Pred. No. 2.4e-269;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRFKDLGEENFKALVLIAPAFQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 1 DAHSEVAHRRFKDLGEENFKALVLIAPAFQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDPNLRLVRREV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDNDPNLRLVRREV 120

QY 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180  
DB 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACLLP 180

QY 181 KLDELDEGKASSAKORLKASLQKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240  
DB 181 KLDELDEGKASSAKORLKASLQKFGGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240

QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAAKOVFLGMFLYEVARRHPDYVSVLLRLAKTYETTLEK 360  
DB 301 DLPSLAADFVESKDVCKNYAAKOVFLGMFLYEVARRHPDYVSVLLRLAKTYETTLEK 360

QY 361 CAADPHCEYAKVDFEPLVEEPONLIKONCELFEOIGEYKFNALLVRYTKVPQVST 420  
DB 361 CAADPHCEYAKVDFEPLVEEPONLIKONCELFEOIGEYKFNALLVRYTKVPQVST 420

QY 421 PTLVEVSRNLGVGSKCKHPEAKMPCAEYDLSVNLQCLVLEHKTVPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGVGSKCKHPEAKMPCAEYDLSVNLQCLVLEHKTVPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDYTPVKFNAETFTFHADICTLSEKEROIKKQALVELVVKHKPKAT 540  
DB 481 LVNRRPCFSALEVDYTPVKFNAETFTFHADICTLSEKEROIKKQALVELVVKHKPKAT 540

QY 541 KEQLKAVMDFFAAFEVKCKKADDETCFAEEGKKLVAAASQAALGL 585  
DB 541 KEQLKAVMDFFAAFEVKCKKADDETCFAEEGKKLVAAASQAALGL 585

## RESULT 7

US-10-319-263-1  
; Sequence 1, Application US/10319263  
; Publication No. US20030180823A1  
; GENERAL INFORMATION:  
; APPLICANT: Bar-Or M.D., David  
; APPLICANT: Lau Ph.D., Edward  
; APPLICANT: Winkler M.D., James V.  
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
; TITLE OF INVENTION: Kits  
; FILE REFERENCE: ISC007  
; CURRENT APPLICATION NUMBER: US/10/319,263  
; CURRENT FILING DATE: 2002-12-13  
; PRIOR APPLICATION NUMBER: 60/115,392  
; PRIOR FILING DATE: 1999-01-11  
; PRIOR APPLICATION NUMBER: 60/102,738  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,926  
; PRIOR FILING DATE: 1998-10-02  
; PRIOR APPLICATION NUMBER: 09/165,581  
; PRIOR FILING DATE: 1998-10-02  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 585  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-319-263-1

```
Query Match      100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDLGEENFKALVLIATFAQYQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 2 DAHKEVAHRRFKDLGEENFKALVLIATFAQYQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRREV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRREV 120
QY 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLIP 180
DB 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLIP 180
QY 181 KLDELDRDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCGHDLLECADRADLAKYICENQDSTSSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLECADRADLAKYICENQDSTSSKLKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADPVESKDVCKNYAAKQVFLGMFLYIYARRHPDYVSWLLRLAKTYETTLK 360
DB 301 DLPSLAADPVESKDVCKNYAAKQVFLGMFLYIYARRHPDYVSWLLRLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLELFEQGEYKFNALLVRYTKVPOVST 420
DB 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLELFEQGEYKFNALLVRYTKVPOVST 420
QY 421 PTLVEVSRNLGVKSGKCKKHPEAKRMPCAEDYLSVWLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGVKSGKCKKHPEAKRMPCAEDYLSVWLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKEROIKKQALVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKEROIKKQALVELVKHKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585
```

RESULT 8

```
US-10-319-263-2
Sequence 2, Application US/103:92e3
Publication No. US20030180820A1
GENERAL INFORMATION:
APPLICANT: Bar-Or M.D., David
APPLICANT: Lau Ph.D., Edward
TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
TITLE OF INVENTION: Kits
FILE REFERENCE: IS007
CURRENT APPLICATION NUMBER: US/10/319,263
CURRENT FILING DATE: 2002-12-13
PRIOR APPLICATION NUMBER: 60/115,392
PRIOR FILING DATE: 1999-01-11
PRIOR APPLICATION NUMBER: 60/102,738
PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: 09/165,926
PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: 09/165,581
PRIOR FILING DATE: 1998-10-02
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 2
LENGTH: 585
TYPE: PRT
```

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: MOD\_RES

LOCATION: (1)..(585)

OTHER INFORMATION: ACETYLATION

US-10-319-263-2

```
Query Match      100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.4e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDLGEENFKALVLIATFAQYQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 2 DAHKEVAHRRFKDLGEENFKALVLIATFAQYQQCFPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRREV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRREV 120
QY 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLIP 180
DB 121 DVMCTAFHNDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLIP 180
QY 181 KLDELDRDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCGHDLLECADRADLAKYICENQDSTSSKLKECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLECADRADLAKYICENQDSTSSKLKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADPVESKDVCKNYAAKQVFLGMFLYIYARRHPDYVSWLLRLAKTYETTLK 360
DB 301 DLPSLAADPVESKDVCKNYAAKQVFLGMFLYIYARRHPDYVSWLLRLAKTYETTLK 360
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLELFEQGEYKFNALLVRYTKVPOVST 420
DB 361 CAADPHCEYAKVDFEFPKPLVEEPQNLKQNCLELFEQGEYKFNALLVRYTKVPOVST 420
QY 421 PTLVEVSRNLGVKSGKCKKHPEAKRMPCAEDYLSVWLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGVKSGKCKKHPEAKRMPCAEDYLSVWLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKEROIKKQALVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETTFHADICTLSEKEROIKKQALVELVKHKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEEGKKLVAAASQAALGL 585
```

RESULT 9

```
US-10-414-469-1
Sequence 1, Application US/104:4469
Publication No. US2003019063A1
GENERAL INFORMATION:
APPLICANT: Bar-Or M.D., David
APPLICANT: Lau Ph.D., Edward
TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
TITLE OF INVENTION: Kits
FILE REFERENCE: IS007
CURRENT APPLICATION NUMBER: US/10/414,469
CURRENT FILING DATE: 2003-04-15
PRIOR APPLICATION NUMBER: 09/806,247
PRIOR FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: PCT/US99/22905
PRIOR FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: 60/115,392
PRIOR FILING DATE: 1999-01-11
PRIOR APPLICATION NUMBER: 60/102,738
PRIOR FILING DATE: 1998-10-02
```

;; PRIOR APPLICATION NUMBER: 09/165,926  
;; PRIOR FILING DATE: 1998-10-02  
;; PRIOR APPLICATION NUMBER: 09/165,581  
;; PRIOR FILING DATE: 1998-10-02  
;; NUMBER OF SEQ ID NOS: 2  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 1  
;; LENGTH: 585  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-414-469-2

Query Match 100.0%; Score 3103; DB 12; Length 585;  
Best Local Similarity 100.0%; Pred. No. 2.4e-269;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 3;  
QY 1 DAHKSEVAHRFKDGLGENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAS 60  
DB 1 DAHKSEVAHRFKDGLGENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAS 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVRPEV 120  
QY 121 DVMTAFPHNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAKAAACLLP 180  
DB 121 DVMTAFPHNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAKAAACLLP 180  
QY 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK 240  
DB 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK 240  
QY 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVSEKDVCKNYAEAKDVFGLGMFLVEYARRHPDYSVWLLRLAKTYETTTLEK 360  
DB 301 DLPSLAADFVSEKDVCKNYAEAKDVFGLGMFLVEYARRHPDYSVWLLRLAKTYETTTLEK 360  
QY 361 CAAADPHCYAKVDFEFPKLVPEEPQNLKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 420  
DB 361 CAAADPHCYAKVDFEFPKLVPEEPQNLKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCCGHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKKCTES 480  
DB 421 PTLVEVSRNLGKVGSKCCGHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKKCTES 480  
QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKHKPKAT 540  
DB 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKHKPKAT 540  
QY 541 KEQLKAVMDDFAAAFVBEKCKCKADDDKTCFAEBEGKLVAAASQAALGL 585  
DB 541 KEQLKAVMDDFAAAFVBEKCKCKADDDKTCFAEBEGKLVAAASQAALGL 585

RESULT 10  
US-10-414-469-2  
;; Sequence 2, Application US/104:4469  
;; Publication No. US20030190691A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Bar-Or M.D., David  
;; APPLICANT: Lau Ph.D., Edward  
;; APPLICANT: Winkler M.D., James V.  
;; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
;; TITLE OF INVENTION: Kits  
;; FILE REFERENCE: ISCO07  
;; CURRENT APPLICATION NUMBER: US/10/414,469  
;; CURRENT FILING DATE: 2003-04-15  
;; PRIOR APPLICATION NUMBER: 09/806,247  
;; PRIOR FILING DATE: 1999-10-01  
;; PRIOR APPLICATION NUMBER: PCT/US99/22905

;; PRIOR FILING DATE: 1999-10-01  
;; PRIOR APPLICATION NUMBER: 60/115,392  
;; PRIOR FILING DATE: 1999-01-11  
;; PRIOR APPLICATION NUMBER: 60/102,738  
;; PRIOR FILING DATE: 1998-10-02  
;; PRIOR APPLICATION NUMBER: 09/165,926  
;; PRIOR FILING DATE: 1998-10-02  
;; PRIOR APPLICATION NUMBER: 09/165,581  
;; PRIOR FILING DATE: 1998-10-02  
;; NUMBER OF SEQ ID NOS: 2  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 2  
;; LENGTH: 585  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: MOD RES  
;; LOCATION: (1)..(585)  
;; OTHER INFORMATION: ACETYLATION  
US-10-414-469-2

Query Match 100.0%; Score 3103; DB 12; Length 585;  
Best Local Similarity 100.0%; Pred. No. 2.4e-269;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKSEVAHRFKDGLGENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAS 60  
DB 1 DAHKSEVAHRFKDGLGENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAS 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVRPEV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVRPEV 120  
QY 121 DVMTAFPHNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAKAAACLLP 180  
DB 121 DVMTAFPHNEETFLKKYLYEIAARRHPYFAPPELLFFAKRYKAAFTCCQAAKAAACLLP 180  
QY 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK 240  
DB 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSORFFKAEFAEVSKLVTDLTK 240  
QY 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 300  
DB 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECKEPLLEKSHCIAEVENDEMPA 300  
QY 301 DLPSLAADFVSEKDVCKNYAEAKDVFGLGMFLVEYARRHPDYSVWLLRLAKTYETTTLEK 360  
DB 301 DLPSLAADFVSEKDVCKNYAEAKDVFGLGMFLVEYARRHPDYSVWLLRLAKTYETTTLEK 360  
QY 361 CAAADPHCYAKVDFEFPKLVPEEPQNLKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 420  
DB 361 CAAADPHCYAKVDFEFPKLVPEEPQNLKQNCLEFQELGEYKFNQALLVRYTKKVPQVST 420  
QY 421 PTLVEVSRNLGKVGSKCCGHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKKCTES 480  
DB 421 PTLVEVSRNLGKVGSKCCGHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKKCTES 480  
QY 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKHKPKAT 540  
DB 481 LVNRRPCFSALEVDITYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKHKPKAT 540  
QY 541 KEQLKAVMDDFAAAFVBEKCKCKADDDKTCFAEBEGKLVAAASQAALGL 585  
DB 541 KEQLKAVMDDFAAAFVBEKCKCKADDDKTCFAEBEGKLVAAASQAALGL 585

RESULT 11  
US-09-932-322-445  
;; Sequence 445, Application US/09932322  
;; Publication No. US20030194743A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Dyax Corp.  
;; APPLICANT: Beltzer, James P.

; APPLICANT: Potter, M. Daniel;  
 ; APPLICANT: Fleming, Tony J.  
 ; APPLICANT: Ladner, Robert Charles  
 ; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (B-lys)  
 ; FILE REFERENCE: DXX-018.1 PCT; DXX-018.1 US  
 ; CURRENT APPLICATION NUMBER: US/09/932.322  
 ; CURRENT FILING DATE: 2001-08-17  
 ; NUMBER OF SEQ ID NOS: 458  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 445  
 ; TYPE: PRT  
 ; ORGANISM: HomoSapiens  
 US-09-932-322-445

Query Match 100.0%; Score 3103; DB 12; Length 585;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-269;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DAHKSEVAHRFKDGLGEEFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE	60
DB	1	DAHKSEVAHRFKDGLGEEFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE	60
QY	61	NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRPEV	120
DB	61	NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRPEV	120
QY	121	DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP	180
DB	121	DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP	180
QY	181	KLDELDEGKASSAKORLKASLOKFGRAFKAAMAVARLSORFFKAEPAEVSUKLVTDLT	240
DB	181	KLDELDEGKASSAKORLKASLOKFGRAFKAAMAVARLSORFFKAEPAEVSUKLVTDLT	240
QY	241	VHTECHGDLLECAADRADLAKYICENQDSISSKKECCCKPLLEKSHCIAEVENDEMPA	300
DB	241	VHTECHGDLLECAADRADLAKYICENQDSISSKKECCCKPLLEKSHCIAEVENDEMPA	300
QY	301	DLPSSAAAFVESKDVCKKYAEAKOVFLGMFLYEAARRHPDYSVLLRLAKTYETTTLEK	360
DB	301	DLPSSAAAFVESKDVCKKYAEAKOVFLGMFLYEAARRHPDYSVLLRLAKTYETTTLEK	360
QY	361	CAAADPHCYAKVDFEFKPLVEEPNLKQNCLEFEQGEYKFNQALVRYTKKVPQVST	420
DB	361	CAAADPHCYAKVDFEFKPLVEEPNLKQNCLEFEQGEYKFNQALVRYTKKVPQVST	420
QY	421	PTLVEVSRNLGKVGSKCKKHPEAKMPCAEEDYLSVNLQCLVLRHEKTPVSDRVTKCTES	480
DB	421	PTLVEVSRNLGKVGSKCKKHPEAKMPCAEEDYLSVNLQCLVLRHEKTPVSDRVTKCTES	480
QY	481	LNNRPPCSALEVDSTYVPKFNABTFTFHADICTLSEKERQIKKOTALVELVGHKPKAT	540
DB	481	LNNRPPCSALEVDSTYVPKFNABTFTFHADICTLSEKERQIKKOTALVELVGHKPKAT	540
QY	541	KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL	585
DB	541	KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL	585

RESULT 12  
 US-10-413-831-1  
 ; Sequence 1, Application US/10413831  
 ; Publication No. US20030194813A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bar-Or M.D., David  
 ; APPLICANT: Lau Ph.D., Edward  
 ; APPLICANT: Winkler M.D., James V.  
 ; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
 ; TITLE OF INVENTION: Kits  
 ; FILE REFERENCE: IS0007  
 ; CURRENT APPLICATION NUMBER: US/10/413,831  
 ; CURRENT FILING DATE: 2003-04-15

; PRIOR APPLICATION NUMBER: US/09/806,247  
 ; PRIOR FILING DATE: 2001-07-16  
 ; PRIOR APPLICATION NUMBER: 60/115,392  
 ; PRIOR FILING DATE: 1999-01-11  
 ; PRIOR APPLICATION NUMBER: 60/102,738  
 ; PRIOR FILING DATE: 1998-10-02  
 ; PRIOR APPLICATION NUMBER: 09/165,926  
 ; PRIOR FILING DATE: 1998-10-02  
 ; PRIOR APPLICATION NUMBER: 09/165,581  
 ; PRIOR FILING DATE: 1998-10-02  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO: 1  
 ; LENGTH: 585  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-413-831-1

QY	1	DAHKSEVAHRFKDGLGEEFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE	60
DB	1	DAHKSEVAHRFKDGLGEEFKALVLIATFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE	60
QY	61	NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRPEV	120
DB	61	NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPRLVLRPEV	120
QY	121	DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP	180
DB	121	DVMCTAFHNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP	180
QY	181	KLDELDEGKASSAKORLKASLOKFGRAFKAAMAVARLSORFFKAEPAEVSUKLVTDLT	240
DB	181	KLDELDEGKASSAKORLKASLOKFGRAFKAAMAVARLSORFFKAEPAEVSUKLVTDLT	240
QY	241	VHTECHGDLLECAADRADLAKYICENQDSISSKKECCCKPLLEKSHCIAEVENDEMPA	300
DB	241	VHTECHGDLLECAADRADLAKYICENQDSISSKKECCCKPLLEKSHCIAEVENDEMPA	300
QY	301	DLPSSAAAFVESKDVCKKYAEAKOVFLGMFLYEAARRHPDYSVLLRLAKTYETTTLEK	360
DB	301	DLPSSAAAFVESKDVCKKYAEAKOVFLGMFLYEAARRHPDYSVLLRLAKTYETTTLEK	360
QY	361	CAAADPHCYAKVDFEFKPLVEEPNLKQNCLEFEQGEYKFNQALVRYTKKVPQVST	420
DB	361	CAAADPHCYAKVDFEFKPLVEEPNLKQNCLEFEQGEYKFNQALVRYTKKVPQVST	420
QY	421	PTLVEVSRNLGKVGSKCKKHPEAKMPCAEEDYLSVNLQCLVLRHEKTPVSDRVTKCTES	480
DB	421	PTLVEVSRNLGKVGSKCKKHPEAKMPCAEEDYLSVNLQCLVLRHEKTPVSDRVTKCTES	480
QY	481	LNNRPPCSALEVDSTYVPKFNABTFTFHADICTLSEKERQIKKOTALVELVGHKPKAT	540
DB	481	LNNRPPCSALEVDSTYVPKFNABTFTFHADICTLSEKERQIKKOTALVELVGHKPKAT	540
QY	541	KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL	585
DB	541	KEQLKAVMDPFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL	585

RESULT 13  
 US-10-413-831-2  
 ; Sequence 2, Application US/10413831  
 ; Publication No. US20030194913A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bar-Or M.D., David  
 ; APPLICANT: Lau Ph.D., Edward  
 ; APPLICANT: Winkler M.D., James V.  
 ; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and  
 ; TITLE OF INVENTION: Kits

FILE REFERENCE: ISCO07  
CURRENT APPLICATION NUMBER: US/10/413,831  
CURRENT FILING DATE: 2003-04-15  
PRIOR APPLICATION NUMBER: US/09/806,247  
PRIOR FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 60/115,392  
PRIOR FILING DATE: 1999-01-11  
PRIOR APPLICATION NUMBER: 60/102,739  
PRIOR FILING DATE: 1998-10-02  
PRIOR APPLICATION NUMBER: 69/165,926  
PRIOR FILING DATE: 1998-10-02  
PRIOR APPLICATION NUMBER: 09/165,581  
PRIOR FILING DATE: 1998-10-02  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 585  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: MOD RES  
LOCATION: (1)-(585)  
OTHER INFORMATION: ACETYLATION  
US-10-413-831-2

Query Match: 100.0%; Score 3103; DB 12; Length 585;  
Best Local Similarity 100.0%; Pred. No. 2.4e-269;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKSEVAHREFKDLGGENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEPAKTCVADESAAE 60  
DB 1 DAHKSEVAHREFKDLGGENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEPAKTCVADESAAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVREPV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVREPV 120  
QY 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLIP 180  
DB 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLIP 180  
QY 181 KLDELDEGKASSAKORLKCASLQKGEAFKAWAVARLSQRPKAEFAEVSKLVTDLT 240  
DB 181 KLDELDEGKASSAKORLKCASLQKGEAFKAWAVARLSQRPKAEFAEVSKLVTDLT 240  
QY 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDKMPA 300  
DB 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDKMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFELYEYARRHPDYSVVLRLAKTYETTLK 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFELYEYARRHPDYSVVLRLAKTYETTLK 360  
QY 361 CAAADPEHCYAKVDFDEPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPOVST 420  
DB 361 CAAADPEHCYAKVDFDEPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPOVST 420  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDYVVPKEFNAETTFHADICTLSEKEROIKKOTALVELVKKPKAT 540  
DB 481 LVNRRPCFSALEVDYVVPKEFNAETTFHADICTLSEKEROIKKOTALVELVKKPKAT 540  
QY 541 KEQLKAVMDFAAFVEKCCCKADDKETCFABEGKLVAAASQAALGL 585  
DB 541 KEQLKAVMDFAAFVEKCCCKADDKETCFABEGKLVAAASQAALGL 585

RESULT 14  
US-10-153-064-5  
Sequence 5, Application US/10/53064

Publication No. US200201428-4A1  
GENERAL INFORMATION:  
APPLICANT: Bell et al.  
TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins  
FILE REFERENCE: PF556  
CURRENT APPLICATION NUMBER: US/10/153,064  
CURRENT FILING DATE: 2002-05-24  
PRIOR APPLICATION NUMBER: 60/293,212  
PRIOR FILING DATE: 2001-05-25  
NUMBER OF SEQ ID NOS: 137  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5  
LENGTH: 585  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-153-064-5  
Query Match: 100.0%; Score 3103; DB 14; Length 585;  
Best Local Similarity 100.0%; Pred. No. 2.4e-269;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DAHKSEVAHREFKDLGGENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEPAKTCVADESAAE 60  
DB 1 DAHKSEVAHREFKDLGGENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEPAKTCVADESAAE 60  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVREPV 120  
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVREPV 120  
QY 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLIP 180  
DB 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLIP 180  
QY 181 KLDELDEGKASSAKORLKCASLQKGEAFKAWAVARLSQRPKAEFAEVSKLVTDLT 240  
DB 181 KLDELDEGKASSAKORLKCASLQKGEAFKAWAVARLSQRPKAEFAEVSKLVTDLT 240  
QY 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDKMPA 300  
DB 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDKMPA 300  
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFELYEYARRHPDYSVVLRLAKTYETTLK 360  
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFELYEYARRHPDYSVVLRLAKTYETTLK 360  
QY 361 CAAADPEHCYAKVDFDEPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPOVST 420  
DB 361 CAAADPEHCYAKVDFDEPKPLVEEPQNLKONCELFQOLGEYKFNALLVRYTKVPOVST 420  
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480  
QY 481 LVNRRPCFSALEVDYVVPKEFNAETTFHADICTLSEKEROIKKOTALVELVKKPKAT 540  
DB 481 LVNRRPCFSALEVDYVVPKEFNAETTFHADICTLSEKEROIKKOTALVELVKKPKAT 540  
QY 541 KEQLKAVMDFAAFVEKCCCKADDKETCFABEGKLVAAASQAALGL 585  
DB 541 KEQLKAVMDFAAFVEKCCCKADDKETCFABEGKLVAAASQAALGL 585  
RESULT 15  
US-09-984-010-7  
Sequence 7, Application US/09984010  
Publication No. US20030104578A1  
GENERAL INFORMATION:  
APPLICANT: Ballance, David James  
TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE  
AND SERUM ALBUMIN  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & JUNNER, LLP

STREET: 1300 I Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/984,010  
FILING DATE: 21-May-2002  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/091,873  
FILING DATE: 25-JUN-1998  
APPLICATION NUMBER: PCT/GB96/03164  
FILING DATE: 19-DEC-1996  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 604 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-984-010-7

Query Match 100.0%; Score 3103; DB 1; Length 604;  
Best Local Similarity 100.0%; Pred. No. 2,5e-269;  
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DAHKSEVAHRPKDGEENFKALVLIAPQYLQCPFDHVKLVNEVTEFAKTCVADES	60
DB	20	DAHKSEVAHRPKDGEENFKALVLIAPQYLQCPFDHVKLVNEVTEFAKTCVADES	79
QY	61	NDCKSLHTLFGDKCTVATLRETYGEMADCCAKOEPERNECFQHKDDNPNLRLVRPEV	120
DB	80	NDCKSLHTLFGDKCTVATLRETYGEMADCCAKOEPERNECFQHKDDNPNLRLVRPEV	139
QY	121	DVNCCTAFHDNEETFLKKYLYEARRHPYFYAPELFFAKRYKAFTCCQAADKAACLLP	180
DB	140	DVNCCTAFHDNEETFLKKYLYEARRHPYFYAPELFFAKRYKAFTCCQAADKAACLLP	199
QY	191	KLDLRDEGKASSAKQLKASLOKGERAFKAWAVARLSQRPFKARFAYSKLVTDLTK	240
DB	200	KLDELRDEGKASSAKQLKASLOKGERAFKAWAVARLSQRPFKARFAYSKLVTDLTK	259
QY	241	VHTECCHGDLLECADDRADLAKYICENQDSISSKLCCKEKPJLEKSHCAEVENDEMPA	300
DB	260	VHTECCHGDLLECADDRADLAKYICENQDSISSKLCCKEKPJLEKSHCAEVENDEMPA	319
QY	301	DLPSLAADFVESKDVCKQNYAEAKOVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK	360
DB	320	DLPSLAADFVESKDVCKQNYAEAKOVFLGMFLYEYARRHPDYSVLLRLAKTYETTLK	379
QY	361	CAAADPHCYAKVDEKFLVEZPQNLKQNCCLFELGEYKFCNALVRYTKKVPQVST	420
DB	380	CAAADPHCYAKVDEKFLVEZPQNLKQNCCLFELGEYKFCNALVRYTKKVPQVST	439
QY	421	PTLVEVSRNLGKSGKCKHPFAKMPCAEDYLSVLNCLCVLHEKTPVSDRVTCKCTES	480
DB	440	PTLVEVSRNLGKSGKCKHPFAKMPCAEDYLSVLNCLCVLHEKTPVSDRVTCKCTES	499
QY	481	LVNRRPCFSALVEDETYPVPEFNAETTFTHADICTLSEKSRQKKQTALVELVAKHFKAT	540
DB	500	LVNRRPCFSALVEDETYPVPEFNAETTFTHADICTLSEKSRQKKQTALVELVAKHFKAT	559
QY	541	KEQLKAVNDDPFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGJ	585
DB	560	KEQLKAVNDDPFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL	604

Search completed: October 27, 2003, 15:31:04  
Job time : 74 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:15:53 : Search time 43 seconds  
(without alignments)  
1308.341 Million cell updates/sec

Title: US-09-833-117-18

Perfect score: 3103

Sequence: 1 DAHKSEVAHFKDLGEENFK.....TCFNECKKLVAASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 3

Maximum DB seq length: 2300000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_76.\*

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	3103	100.0	609	1	ABHUS		serum albumin prec
2	2942	94.8	600	2	A47391		serum albumin prec
3	2620	84.4	608	2	S57632		serum albumin prec
4	2475.5	79.8	607	1	ABHOS		serum albumin prec
5	2446.5	78.8	607	1	ABHOS		serum albumin prec
6	2432.5	78.4	607	1	ABHOS		serum albumin prec
7	2426	78.2	608	1	ABHOS		serum albumin prec
8	2411.5	77.7	605	1	ABHOS		serum albumin prec
9	2387	76.9	609	2	CS5838		albumin - Xongolia
10	1861	60.3	453	2	A51339		serum albumin - mo
11	1557.5	50.2	615	1	ABHOS		serum albumin prec
12	1253.5	40.4	609	2	UC4258		alpha-fetoprotein
13	1249.5	40.3	609	1	FPHU		alpha-fetoprotein
14	1242.5	40.0	603	1	FPHU		alpha-fetoprotein
15	1205	38.8	607	1	ABHOS		74K albumin prec
16	1181.5	38.1	265	2	A49886		albumin - dog liver
17	1175.5	37.9	608	2	ABHOS		68K serum albumin
18	1084	34.9	605	1	FPHU		alpha-fetoprotein
19	1067	34.4	611	1	FPHU		alpha-fetoprotein
20	1055	34.0	599	1	A54906		albumin precursor
21	928.5	29.9	614	2	S59517		serum albumin prec
22	928	29.9	608	2	A53195		albumin precursor
23	747.5	24.1	608	1	ABHOS		serum albumin 2 pr
24	742.5	23.9	608	1	ABHOS		serum albumin 2 pr
25	699	22.5	382	2	A37253		serum albumin - bu
26	440.5	14.2	423	1	S27941		serum albumin - se
27	386	12.4	474	1	VHVD		vitamin D-binding
28	385	12.4	476	1	VHVD		vitamin D-binding
29	372	12.0	472	1	A35327		vitamin D-binding

RESULT 1

ABHUS

serum albumin precursor [validated] - human

N;Alternate names: preproalbumin

N;Contains: kinetensin

C;Species: Homo sapiens (man)

C;Date: 29-Jul-1981 #sequence, revision 31-Jan-1997 #text change 17-Mar-2000

C;Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S56422;

R;Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; See

Nucleic Acids Res. 9, 6103-6114, 1981

A;Title: The sequence of human serum albumin cDNA and its expression in Escherichia

A;Reference number: A93743; MUID:82081892; PMID:6171778

A;Accession: A93743

A;Molecule type: mRNA

A;Residues: 1-419, 'K', 421-609 <LAW>

A;Cross-references: EMBL:V00495; GB:J00078; GB:J00133; GB:J00133; NID:G28591; PIDN:CF

R;Dugaiczkyk, A.; Law, S.W.; Dennison, O.E.

Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982

A;Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.

A;Reference number: A93936; MUID:82105994; PMID:6275391

A;Accession: A93936

A;Molecule type: mRNA

A;Residues: 1-120, 'G', 122-609 <DUG>

A;Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; P.D:G28590

R;Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.

J. Biol. Chem. 261, 3244-3251, 1986

A;Title: The human albumin gene. Characterization of the 5' and 3' flanking regions

A;Reference number: I39427; MUID:86140099; PMID:2419329

A;Accession: I39427

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-26 <URA>

A;Cross-references: GB:M13075; NID:G178330; PIDN:AAAS1688.1; PID:G553173

R;Watkins, S.; Radison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.

Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994

A;Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian fami

A;Reference number: I59286; MUID:94181575; PMID:8134387

A;Accession: I59286

A;Status: translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 282-290, 'KSRFPLC', <WAT>

A;Cross-references: GB:S69192; NID:G546032; PIDN:AAB30282.1; PID:G546033

A;Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia

R;Madison, G.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putra

Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994

A;Title: Genetic variants of human serum albumin in Italy: point mutants and a carbox

A;Reference number: I59313; MUID:94294404; PMID:8022807

A;Accession: I59313

A;Status: translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Cross-references: GB:S70799; NID:G547231; PIDN:AAB31177.1; P.D:G547232

cag island protein  
cag pathogenicity  
calcium-binding pr  
hypothetical prote  
major surface glyco  
cell surface glyco  
hypothetical prote  
sperm tail-specific  
kinesin homolog F2  
major surface glyco  
myosin heavy chain  
cell-cycle-depende  
glycoprotein A - P  
hypothetical prote  
granin - human  
embryonic muscle m

A:Note: this frame-shift variant is designated albumin Bazzano; four additional variants  
R:Menaya, J.; Parrilla, R.; Ayuso, M.S.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: G08292  
A:Accession: G01747  
A:Status: translated from GB/EMBL/DBSJ  
A:Molecule type: mRNA  
A:Residues: 1-120, 'G' 122-455 <MEN>  
A:Cross-references: EMBL:U22961; NID:G763428; PID:AAA64922.1; PID:G763431  
R:Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.  
Biochem. J. 308, 321-325, 1995  
A:Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2  
A:Reference number: S55314; MUID:95275251; PMID:7755581  
A:Accession: S55314  
A:Molecule type: protein  
A:Residues: 19-27 <LED>  
R:Meiloun, B.; Moravsek, L.; Kostka, V.  
FEBS Lett. 58, 134-137, 1975  
A:Title: Complete amino acid sequence of human serum albumin.  
A:Reference number: A9420; MUID:76187907; PMID:1225573  
A:Accession: A9420  
A:Molecule type: protein  
A:Residues: 25-117, 'EQ' 120-154, 'Q' 156-193, 'E' 195-387, 'H' 389-390, 'Y' 392-393, 'A' 395-  
R:Roehr, U.; Spitteler, G.; Tripler, E.  
Justus Liebig's Ann. Chem. 9, 881-884, 1988  
A:Title: Isolation and structure elucidation of middle-molecular weight peptides from u  
A:Reference number: S06422  
A:Note: this paper is in German, with an English abstract  
A:Accession: S06422  
A:Molecule type: protein  
A:Residues: 25-48 <ROE>  
R:Pinch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.  
Arch. Biochem. Biophys. 305, 595-599, 1993  
A:Title: Mass spectrometric identification of modifications to human serum albumin treat  
A:Reference number: S38982; MUID:93394321; PMID:8373198  
A:Accession: S38982  
A:Molecule type: protein  
A:Residues: 45-67; 141-160; 311-337; 469-490; 570-581 <FIN>  
R:Kausler, E.; Spitteler, G.  
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991  
A:Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mitte  
A:Reference number: S17599; MUID:92126241; PMID:1772598  
A:Accession: S17599  
A:Molecule type: protein  
A:Residues: 25-54; 354-357; 431-447 <KAU>  
A:Note: 49-Leu was also found  
R:Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A:Title: Structures of histamine-releasing peptides formed by the action of acid proteas  
A:Reference number: A45800; MUID:89341456; PMID:2474609  
A:Accession: A45800  
A:Molecule type: protein  
A:Residues: 166-173 <CAR>  
R:Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa  
Biochem. Biophys. Res. Commun. 136, 983-988, 1986  
A:Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre  
A:Reference number: A03239; MUID:86242180; PMID:3087352  
A:Accession: A03239  
A:Molecule type: protein  
A:Residues: 166-173, 'L' <XOG>  
R:Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S  
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990  
A:Title: Mutations in genetic variants of human serum albumin found in Italy.  
A:Reference number: A38255; MUID:91062352; PMID:2247440  
A:Accession: C38255  
A:Molecule type: protein  
A:Residues: 76-111 <GAL1>  
A:Accession: B38255  
A:Molecule type: protein  
A:Residues: 82-105, 'K' 107-110 <GAL2>  
A:Note: this variant is designated albumin Vibo Valentia  
A:Accession: A38255  
A:Molecule type: protein  
A:Residues: 76-83, 'K' 85-156 <GAL3>  
A:Note: this variant is designated albumin Torino  
R:Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.  
Eur. J. Biochem. 214, 437-444, 1993  
A:Title: The structural characterization and bilirubin-binding properties of albumin  
A:Reference number: S33298; MUID:93292504; PMID:8513793  
A:Accession: S33298  
A:Molecule type: protein  
A:Residues: 255-263, 'E' 265-281 <MIN1>  
A:Note: this variant is designated albumin Herborn  
R:Minchiotti, L.; Galliano, M.; Scoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Po  
Biochim. Biophys. Acta 1119, 232-238, 1992  
A:Title: Two albumins with identical electrophoretic mobility are produced by dif  
A:Reference number: S21078; MUID:92190239; PMID:1347703  
A:Accession: S21078  
A:Molecule type: protein  
A:Residues: 354-356, 'K' 358-378 <MIN2>  
A:Note: this variant is designated albumin Sondrio; another variant Paris-2 is repor  
R:He, X.M.; Carter, D.C.  
Nature 358, 209-215, 1992  
A:Title: Atomic structure and chemistry of human serum albumin.  
A:Reference number: A46756; MUID:92334427; PMID:1630489  
A:Contents: annotation; x-ray crystallography, 2.8 angstroms  
R:Brown, J.R.; Stockley, P.; Behrens, P.O.  
In The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23  
A:Reference number: A94442  
A:Contents: annotation; three-dimensional structure and disulfide bonds  
R:Sabat, M.A.; Stockbauer, P.; Moravsek, L.; Meiloun, B.  
Collect. Czech. Chem. Commun. 42, 564-579, 1977  
A:Title: Disulfide bonds in human serum albumin.  
A:Reference number: A90930  
A:Contents: annotation; disulfide bonds  
R:Jacobsen, C.  
Biochem. J. 17, 453-459, 1978  
A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity bind  
A:Reference number: A90299; MUID:78186630; PMID:656055  
A:Contents: annotation; bilirubin-binding site  
R:Peters, T.; Reed, R.G.  
In Albumin: Structure, Biosynthesis, Function, Peters, J., and Scholm, I., eds., 11  
A:Title: Serum albumin: conformation and active sites.  
A:Reference number: A94408  
A:Contents: annotation; binding sites  
R:Harper, M.E.; Dugalczyk, A.  
Am. J. Hum. Genet. 35, 565-572, 1983  
A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein g  
A:Reference number: A90028; MUID:81279982; PMID:6192711  
A:Contents: annotation; gene position  
R:Walker, J.E.  
FEBS Lett. 66, 173-175, 1976  
A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic a  
A:Reference number: A46755; MUID:76257808; PMID:955075  
A:Contents: annotation  
A:Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic  
R:Donney, J.P.; Fonda, M.L.; Feldhoff, R.C.  
FEBS Lett. 298, 266-268, 1992  
A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-pho  
A:Reference number: A56294; MUID:92183881; PMID:1544460  
A:Contents: annotation  
A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described;  
ase activity  
C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthe  
lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (w  
C:Comment: A large number of variants of human serum albumin have been described.  
C:Genetics:  
A:Gene: GDB:ALB  
A:Cross-references: GDB:118990; OMIM:103600  
A:Map position: 4q11-4q13  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; py  
F.1-18/Domain: signal sequence #status predicted <SIG>  
F.19-24/Domain: propeptide #status experimental <PRO>  
F.25-609/Product: serum albumin #status experimental <MPT>  
F.29-202/Domain: serum albumin repeat homology <SA1>

F:166-174/Product: kinotensin serum experimental <KIP>  
F:221-394/Domain: serum albumin repeat homology <SA2>  
F:413-592/Domain: serum albumin repeat homology <SA2>  
F:27/Binding site: copper (His) #status predicted  
F:77-86,95-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-395,394-393,4  
F:214/Binding site: pyridoxal phosphate (lys) (covalent) #status experimental

Query Match 100.0%; Score 3.03; DB 1; Length 609;  
Best Local Similarity 100.0%; Pred. No. 9,28-198; Mismatches 0; Indels 0; Gaps 0;  
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDLGSENFKALVLIARFYQYQCCPEEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 25 DAHKEVAHRRFKDLGSENFKALVLIARFYQYQCCPEEDHVKLVNEVTEFAKTCVADESAAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLPLVRREV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLPLVRREV 144  
QY 121 DVMCTAFHONETETLKKLYEYIARRHPYFAPPELLFFPAARYKAAFTBCCQAADKAACLLP 180  
DB 145 DVMCTAFHONETETLKKLYEYIARRHPYFAPPELLFFPAARYKAAFTBCCQAADKAACLLP 204  
QY 181 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEYKLVTDITK 240  
DB 205 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEYKLVTDITK 264  
QY 241 VHTECCHGDLLECAADRADLAKYICENQDSISSKLEKCECPLEKSHCIAFVENDEMPA 300  
DB 265 VHTECCHGDLLECAADRADLAKYICENQDSISSKLEKCECPLEKSHCIAFVENDEMPA 324  
QY 301 DLPSLAANDVESKQVCKNYAEAKQVFLGMFLYEAARRHPDYSVLLILAKTYETTLK 360  
DB 325 DLPSLAANDVESKQVCKNYAEAKQVFLGMFLYEAARRHPDYSVLLILAKTYETTLK 384  
QY 361 CAADDPHECYAKVDFEFPKPLVEBPQNLKQNCLEPEQLGEYKFNALLVRYTKVPQVST 420  
DB 385 CAADDPHECYAKVDFEFPKPLVEBPQNLKQNCLEPEQLGEYKFNALLVRYTKVPQVST 444  
QY 421 PTLVEVSRNLGVKGVKCKKHPKAPCAEDYLSVNLQVLVHEKTPVSDRVTKCCTES 480  
DB 445 PTLVEVSRNLGVKGVKCKKHPKAPCAEDYLSVNLQVLVHEKTPVSDRVTKCCTES 504  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKQATLVELVGHKPKAT 540  
DB 505 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKQATLVELVGHKPKAT 564  
QY 541 KEQLKAVMDQFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 585  
DB 565 KEQLKAVMDQFAAFVEKCKCKADKCTCFABEGKKLVAASQAALGL 609

RESULT 2  
A47391  
serum albumin precursor - rhesus macaque  
C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #ext\_change 20-Aug-1995  
C:Accession: A47391  
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, B.; Smith, D.G.; Dwulet, J.; Putnam, F.  
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993  
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in biliary  
A:Reference number: A47391; MUID:93211971; PMID:8460152  
A:Contents: B/B homozygote  
A:Accession: A47391  
A:Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 1-600 <WAT>  
A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36306.1; PID:G342295  
A:Experimental source: liver  
A:Note: sequence extracted from NCBI backbone (NCBI:128282, NCBI:P:128281)  
C:Superfamily: serum albumin; serum albumin repeat homology  
F:21-194/Domain: serum albumin repeat homology <SA1>  
F:213-386/Domain: serum albumin repeat homology <SA2>

F:405-584/Domain: serum albumin repeat homology <SA3>  
Query Match 94.8%; Score 2942; DB 2; Length 600;  
Best Local Similarity 93.5%; Pred. No. 48-187;  
Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHKEVAHRRFKDLGSENFKALVLIARFYQYQCCPEEDHVKLVNEVTEFAKTCVADESAAE 60  
DB 17 DTHKSEVAHRRFKDLGSENFKALVLIARFYQYQCCPEEDHVKLVNEVTEFAKTCVADESAAE 76  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLPLVRREV 120  
DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPPLPLVRREV 136  
QY 121 DVMCTAFHONETETLKKLYEYIARRHPYFAPPELLFFPAARYKAAFTBCCQAADKAACLLP 180  
DB 137 DVMCTAFHONETETLKKLYEYIARRHPYFAPPELLFFPAARYKAAFTBCCQAADKAACLLP 196  
QY 181 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEYKLVTDITK 240  
DB 197 KLDELDEGKASSAKQRLKASLQKFGGERAFKAWAVARLSQRPFAEFAEYKLVTDITK 256  
QY 241 VHTECCHGDLLECAADRADLAKYICENQDSISSKLEKCECPLEKSHCIAFVENDEMPA 300  
DB 257 VHTECCHGDLLECAADRADLAKYICENQDSISSKLEKCECPLEKSHCIAFVENDEMPA 316  
QY 301 DLPSLAANDVESKQVCKNYAEAKQVFLGMFLYEAARRHPDYSVLLILAKTYETTLK 360  
DB 317 DLPSLAANDVESKQVCKNYAEAKQVFLGMFLYEAARRHPDYSVLLILAKTYETTLK 376  
QY 361 CAADDPHECYAKVDFEFPKPLVEBPQNLKQNCLEPEQLGEYKFNALLVRYTKVPQVST 420  
DB 377 CAADDPHECYAKVDFEFPKPLVEBPQNLKQNCLEPEQLGEYKFNALLVRYTKVPQVST 436  
QY 421 PTLVEVSRNLGVKGVKCKKHPKAPCAEDYLSVNLQVLVHEKTPVSDRVTKCCTES 480  
DB 437 PTLVEVSRNLGVKGVKCKKHPKAPCAEDYLSVNLQVLVHEKTPVSDRVTKCCTES 496  
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKQATLVELVGHKPKAT 540  
DB 497 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKEROIKKQATLVELVGHKPKAT 556  
QY 541 KEQLKAVMDQFAAFVEKCKCKADKCTCFABEGKKLVAASQAAL 583  
DB 557 KEQLKAVMDQFAAFVEKCKCKADKCTCFABEGKKLVAASQAAL 599

RESULT 3  
S57632  
serum albumin precursor - cat  
C:Species: Felis silvestris catus (domestic cat)  
C:Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #ext\_change 20-Aug-1999  
C:Accession: J04660; S57632  
R:Hiiger, C.; Grigioni, F.; Hentges, F.  
Gene 289, 295-296, 1996  
A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.  
A:Reference number: J04660; MUID:96194824; PMID:8647463  
A:Accession: J04660  
A:Molecule type: mRNA  
A:Residues: 1-608 <H12>  
A:Cross-references: EMBL:X84842; NID:G886484; PIDN:CAA59279.1; PID:G886485  
A:Experimental source: liver  
A:Comment: This protein is the major protein component in plasma. It functions as a  
e in has 35 conserved cysteine residues.  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: liver; plasma  
F:1-18/Domain: signal sequence #status predicted <S1>  
F:19-24/Domain: propeptide #status predicted <PRP>  
F:25-608/Product: serum albumin #status predicted <WAT>  
F:25-202/Domain: serum albumin repeat homology <SA1>  
F:221-394/Domain: serum albumin repeat homology <SA2>  
F:413-592/Domain: serum albumin repeat homology <SA3>

```
Query Match      84.4%; Score 2623; DB 2; Length 608;
Best Local Similarity 82.0%; Pred. No. 8, 2e-166;
Matches 478; Conservative 32; Mismatches 53; Indels 0; Gaps 0;

QY 1 DAHKEVAHFKDGLGEENFKALVLIAPQYQQCPFECHVKLVNVEYFAKTCVADESAE 60
DB 25 EAHQSEIAHFNDELGEHFRGVLVAFSQYQQCPFECHVKLVNVEYFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQHKDGNPNLPLRVPEV 120
DB 85 NCEKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQHKDGNPNLPLRVPEV 144
QY 121 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLP 180
DB 145 DAMCTAFHNEEQFLGKYLYEIAARRHPYFYAPPELLFYAEYKGVFTTECCQAADKAACLP 204
QY 181 KLDELDEGKASSAKORLKCASLOKQGERAFKAWAVARLSORFPKAEFAEVSCLVDTLK 240
DB 205 KVDALREKVLASSAKERLKCASLOKQGERAFKAWAVARLSORFPKAEFAEVSCLVDTLK 264
QY 241 VHTCCCHGDLLECAADDRADLAKYICENQDSISSKLECCCKPFLLEKSHCIAEVENDEMPA 300
DB 265 IHKECCCHGDLLECAADDRADLAKYICENQDSISSKLECCCKPFLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADPFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYVSVLLRLAKTYETITLKC 360
DB 325 DLPPLAVDEVEDKVCNKYQAEAKDVFLGTLFYYSRRHPDYVSVLLRLAKTYETITLKC 384
QY 361 CAAADPHCEYAKVDFEKLVEEPQNLIKONCELFQOLGEYFQNALVRYTKVPQVST 420
DB 385 CATDDPPACVAVDFEKLVEEPQNLIKONCELFQOLGEYFQNALVRYTKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNLQJCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNLQJCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVEYVKEKPKAT 540
DB 505 LVNRRCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVEYVKEKPKAT 564
QY 541 KEQLKAWMODFAAFVCKCKADDDKETCFABEGKLVAAASQAL 583
DB 565 EEQLKATVMGDFGSDVKCCAAEDKEACFAEEGPKLVAAASQAL 607

RESULT 4
ABOS
serum albumin precursor - horse
C:Species: Equus caballus (domestic horse)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S34053
R:Ho, J.X.; Holowachuk, E.W.; Norton, E.C.; Twigg, F.D.; Carter, D.C.
Eur. J. Biochem. 215, 205-212, 1993
A:Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
A:Reference number: S34053; MUID:93345495; PMID:8344282
A:Accession: S34053
A:Molecule type: mRNA
A:Residues: 1-607 <HOA>
A:Cross-references: GB:X74045; NID:G399671; PID:CAAS2194.1; PID:G399672
A:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <YAT>
F:29-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:2/Binding site: copper (His) #status predicted
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F:263/Binding site: bilirubin (Iys) #status predicted
```

```
Query Match      79.8%; Score 2475.5; DB 1; Length 607;
Best Local Similarity 76.3%; Pred. No. 3e-156;
Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1;

QY 1 DAHKEVAHFKDGLGEENFKALVLIAPQYQQCPFECHVKLVNVEYFAKTCVADESAE 60
DB 25 DTHKSEIAHFNDELGEHFRGVLVAFSQYQQCPFECHVKLVNVEYFAKTCVADESAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQHKDGNPNLPLRVPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQHKDGNPNLPLRVPEV 143
QY 121 DVMCTAFHNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLP 180
DB 144 DAOCAAFQEDPDKFLGKYLYEVARRHPYFYGPPELLPHAEYKADFTTECCQAADKAACLP 203
QY 181 KLDELDEGKASSAKORLKCASLOKQGERAFKAWAVARLSORFPKAEFAEVSCLVDTLK 240
DB 204 KLDALKERILLSSAKERLKCASSFQNGERAFKAWAVARLSORFPKAEFAEVSCLVDTLK 263
QY 241 VHTCCCHGDLLECAADDRADLAKYICENQDSISSKLECCCKPFLLEKSHCIAEVENDEMPA 300
DB 264 VKHCCCHGDLLECAADDRADLAKYICENQDSISSKLECCCKPFLLEKSHCIAEVENDEMPA 323
QY 301 DLPSLAADPFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYVSVLLRLAKTYETITLKC 360
DB 324 DLPALAADFAEDKEICXHYKDAKDVFLGTLFYYSRRHPDYVSVLLRLAKTYETITLKC 383
QY 361 CAAADPHCEYAKVDFEKLVEEPQNLIKONCELFQOLGEYFQNALVRYTKVPQVST 420
DB 384 CAEADPPACVAVDFEKLVEEPQNLIKONCELFQOLGEYFQNALVRYTKVPQVST 443
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNLQJCVLHEKTPVSDRVTKCCTES 480
DB 444 PTLVEIGRTJGVKVSRCCKLPSESERLPCSENHIALALNRLCVLHEKTPVSEKTKCCTDS 503
QY 481 LVNRRCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVEYVKEKPKAT 540
DB 504 LAERRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATLVEYVKEKPKAT 563
QY 541 KEQLKAWMODFAAFVCKCKADDDKETCFABEGKLVAAASQAL 583
DB 564 KEQLKATVJGNFSAPFAKCCGREDKEACFAEEGPKLVAAASQAL 606

RESULT 5
ABOS
serum albumin precursor [validated] - bovine
N:Alternate names: 67K protein; preproalbumin
C:Species: Bos primigenius taurus (cattle)
C>Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000
C:Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A31458;
R:Holowachuk, E.W.; Stoltzberg, J.K.; Reed, R.G.; Peters Jr., T.
Submitted to the EMBL Data Library, August 1991
A:Description: Bovine serum albumin: cDNA sequence and expression.
A:Reference number: A38885
A:Accession: A38885
A:Molecule type: mRNA
A:Residues: 1-607 <HOA>
A:Cross-references: EMBL:M73215
R:Harayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.
Biochem. Biophys. Res. Commun. 173, 639-646, 1990
R:MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.
Eur. J. Biochem. 98, 477-485, 1979
A:Title: Rapid confirmation and revision of the primary structure of bovine serum a)
A:Reference number: A36401; MUID:91083649; PMID:2260975
A:Accession: A36401
A:Molecule type: protein
A:Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>
R:MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.
Eur. J. Biochem. 98, 477-485, 1979
A:Title: Biosynthesis of bovine plasma proteins in a cell-free system.
A:Reference number: A91258; MUID:80024278; PMID:488109
A:Accession: A91258
A:Molecule type: protein
```

A:Residues: 1-32 <YAG>  
R.Hsieh, J.C.; Lin, F.P.; Tam, M.F.  
Anal. Biochem. 170, 1-9, 1988  
A:Title: Electrophoretic transfer of an analytical isoelectrofocusing gel  
A:Reference number: A60808; MUID:88267456; PMID:3389500  
A:Accession: B60808  
A:Molecule type: protein  
A:Residues: 25-41 <HSI>  
R.Strawick, E.; Glanville, M.J.  
Eur. J. Biochem. 191, 47-56, 1990  
A:Title: Toxin 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albumin  
A:Reference number: S13780; MUID:90336641; PMID:2379503  
A:Accession: S10780  
A:Molecule type: protein  
A:Residues: 25-41, 43-57, 59-64 <STR>  
R.Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1694, 1989  
A:Title: Structures of histamine-releasing peptides formed by the action of acid proteases  
A:Reference number: A45800; MUID:89341406; PMID:2474609  
A:Accession: D45800  
A:Molecule type: protein  
A:Residues: 163-172 <CAR>  
R.Carraway, R.E.; Mitra, S.P.; Cochrane, D.E.  
J. Biol. Chem. 262, 5968-5973, 1987  
A:Title: Structure of a biologically active neurotensin-related peptide obtained from pe  
A:Reference number: A26693; MUID:87194805; PMID:243711  
A:Accession: A26693  
A:Molecule type: protein  
A:Residues: 165-172, 174 <CA2>  
R.Reed, R.G.; Putnam, F.W.; Peters Jr., T.  
Biochem. J. 191, 867-868, 1980  
A:Title: Sequence of residues 400-403 of bovine serum albumin.  
A:Reference number: A90309; MUID:82023364; PMID:7283978  
A:Accession: A90309  
A:Molecule type: protein  
A:Residues: 402-433 <R3E>  
R.Brown, J.R.  
Fed. Proc. 34, 591, 1975  
A:Title: Structure of bovine serum albumin.  
A:Reference number: A94551  
A:Accession: A94551  
A:Molecule type: protein  
A:Residues: 190-195 <BR2>  
R.Brown, J.R.  
Fed. Proc. 33, 1389, 1974  
A:Reference number: A91457  
A:Contents: annotation; disulfide bonds  
R.Werlen, R.C.; Offord, R.E.; Rose, K.  
Biochem. J. 302, 907-911, 1994  
A:Title: Preparation and characterization of novel substrates of insulin proteinase (EC  
A:Reference number: S55232; MUID:95031935; PMID:7945219  
A:Accession: S55232  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 529-536, 569-572 <WER>  
C:Keywords: carrier protein; serum albumin; repeat homology  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-607/Product: serum albumin #status experimental <MPT>  
F:29-201/Domain: serum albumin repeat homology <SA1>  
F:220-393/Domain: serum albumin repeat homology <SA2>  
F:412-591/Domain: serum albumin repeat homology <SA3>  
F:27/Binding site: copper (His) #status predicted  
F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392, 4

Query Match

78.8%; Score 2446.5; DB 1; Length 607;

Best Local Similarity 75.6%; Pred. No. 2.5e-154;  
Matches 441; Conservative 71; Mismatches 70; Indels 1; Gaps 1;  
Qy 1 DAHSEVAHFRKDLGEENFKALVLIAFACLYLOCCPFEDHVKLVREVTFAKTCVADESAE 60  
Db 25 DTHKSEIAHFRKDLGEBOPKGLVLIAFSCYLOCCPFDEHVKLVNELTEFAKTCVADESHA 84  
Qy 61 NDKSLHTLFGSKLCTVATLRETYGEMADCCAKQPERNECFLOHKKDNDNLPRLVRREV 120  
Db 85 GGEKSLHTLFGDLCKVASLRETYGEMADCCQPERNECFSLKSDSPDLPKL-KEDP 143  
Qy 121 DVNCTAFHNEETFLKVKYVEIARRHPYFAYZELLFFAKRYKAAFTCCCAADKACILP 180  
Db 144 NTLCDPEFKADEKKFWCKYLYEIAARRHPYFAYZELLFYANKYNGVFQDCCOAEKGCILP 203  
Qy 181 KUDELDEGKASAKORLKASLQKFGERAFKAWARLSQFPKAEFAVSKLVTDLTK 240  
Db 204 KIETREKVLASSAROLKASIQKFGERALKAMSVARLSQFKPAEFVEVTKLVTDLTK 263  
Qy 241 VHTCCGGLLECAODRADLAKYICENODSISKKCECEKPLLEKSHCIAVENDEMPA 300  
Db 264 VHKCCGGLLECAODRADLAKYICONDITSSKKECCOKPLLEKSHCIAEVEKDA-PE 323  
Qy 301 DLPSLAADVESKDVCKNYAEAKDFLGMFLYVEYARRHPDYVLLLR-LAKTYETTLK 360  
Db 324 NUPPLTADFAEDKDVCKNYQEAQAFLOGSFVYEYSRRHPZYAVSVLLR-LAKEYEATLEEC 383  
Qy 361 CAAADPHECVAKVDFEKLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKVQVST 420  
Db 384 CAKDDPHACYSTVDFKLHVLDEPQNLKQNCDFEKLGEYKFNALLVRYTRKQVST 443  
Qy 421 PTLVEVSRNLGKVGSKCKKHPKAPCAEDVSVVNLGLCVLHEKTPVSDRVTKCCTES 480  
Db 444 PTLVEVSRNLGKVGTRCCTKPESEMPCTEDVLSILNLRLCVLHEKTPVSEKTKCCTES 503  
Qy 481 LVNRPCFSALEVDVTPKPFNAETFTPHADICTLSEKERQIKKQATLVEVVKHKPKAT 540  
Db 504 LVNRPCFSALETPDEYVPKAFDEKLFTPHADICTLPTDEKQIKKQATLVELLKHKPKAT 563  
Qy 541 KEOLKAVMCDFAFVEKCKCKADKDETCFAEEGKLVVAASQAL 583  
Db 564 EQLKTMENFVAFVDKCAADKACFAVEGPKLVVSTQATL 606  
RESULT 6  
ABSSS  
serum albumin precursor - sheep  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C:Accession: S06936  
R:Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.  
Nucleic Acids Res. 17, 10495, 1989  
A:Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.  
A:Reference number: S06936; MUID:90098888; PMID:2602160  
A:Accession: S06936  
A:Molecule type: mRNA  
A:Residues: 1-607 <BRO>  
A:Cross-references: EMBL:X17055; NID:G1386; PID:CA434903.1; PID:G1387  
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper  
teroid hormones (weak bonds with these hormones promote their transfer across the men  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-24/Domain: propeptide #status predicted <PRO>  
F:25-607/Product: serum albumin #status predicted <MAT>  
F:29-201/Domain: serum albumin repeat homology <SA1>  
F:220-393/Domain: serum albumin repeat homology <SA2>  
F:412-591/Domain: serum albumin repeat homology <SA3>  
F:27/Binding site: copper (His) #status predicted  
F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-393  
F:263/Binding site: bilirubin (Lys) #status predicted

Query Match

78.4%; Score 2432.5; DB 1; Length 607;

```

Best Local Similarity 75.0%; Pred. No. 2.1e-153;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

QY 1 DAHSEVAHRRFKOLGEENPKALVLIAPACYQQCPFDHVKLVNEVTEFAKTCVADSEAE 60
DB 25 DTHKSEIAHRRNDLGEENFQGLVLIAPSYQYQQCPFDHVKLVNEXTEFAKTCVADSEAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVLRPEV 120
DB 85 CQDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVLRPEV 143
QY 121 DVMTAFHNDNEETPLKKYLYEIAIRRHYPFYAPPELLFFAKYKAAFTSCQAAADKACLLP 180
DB 144 DTLCAEPKADKKFKYLYEIAIRRHYPFYAPPELLFFAKYKAAFTSCQAAADKACLLP 203
QY 131 KDELDRDEGKASSAKORLKCASLQKQFGERAFKAWAVARLSQRFKAEFAEYSKLVTLTK 240
DB 204 KIDAMREKVLASSAKORLKCASLQKQFGERAFKAWAVARLSQRFKAEFAEYSKLVTLTK 263
QY 241 VHTCCCHGDLLECCADDRADLAKYICENODSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 264 VHKCECHGDLLECCADDRADLAKYICENODSISSKLEKCEKPLLEKSHCIAEVENDEMPA 323
QY 301 DLPSLAADPFVESKDVCKYKAAKQVFLGMFLYEVARRHPDYSVVLRLAKTYETTLTK 360
DB 324 NPLPLTAQFAEDKEVKCKYKAAKQVFLGMFLYEVARRHPDYSVVLRLAKTYETTLTK 383
QY 361 CAADPHCEYKAVDFEKPVLVEEPQNLKQNCLEFEQOLGEYKFNQALLVRYTKKQPOVST 420
DB 384 CAEDPHCEYKAVDFEKPVLVEEPQNLKQNCLEFEQOLGEYKFNQALLVRYTKKQPOVST 443
QY 421 PLVEYSRNLGKVGKCKKPEAKMPCAEDYLSVLNQCVLREKTPVSRVTKCCTES 480
DB 444 PLVEYSRNLGKVGKCKKPEAKMPCAEDYLSVLNQCVLREKTPVSRVTKCCTES 503
QY 481 LVNRRPCFSALEVDYVYPKFNNAETFTTHADICTLSEKERIKQKQALVELVKKKPKAT 540
DB 504 LVNRRPCFSALEVDYVYPKFNNAETFTTHADICTLSEKERIKQKQALVELVKKKPKAT 563
QY 541 KEQLKAVMDDFAAVEKCKCKADQKCTCFAECKGLVAASQAL 583
DB 564 DEQLKTWENFAVDFKCCAADQKCGFVVBGPKLVASTQAL 606

RESULT 7
ABR'S
serum albumin precursor - rat
N:Alternate names: prealbumin
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-May-1979 #sequence revision 31-May-1979 #text change 22-Jun-1999
C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R:Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981.
A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A:Reference number: A93872; MUID:81223722; PMID:7317712
A:Accession: A93872
A:Molecule type: mRNA
A:Residues: 1-608 <S>
A:Cross-references: GB:J00698; NID:955627; PIDN:CAA24532.1; PID:955628
R:Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rockey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Anal.
A:Reference number: A92211; MUID:77249657; PMID:893447
A:Note: Cleavages during protein maturation
A:Accession: A92211
A:Molecule type: protein
A:Residues: 1-38 <S>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A:Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
A:Reference number: A91946; MUID:78109429; PMID:564345
A:Accession: A91946
A:Molecule type: protein

```

A:Residues: 25-222 <IS1>  
R:Isemura, S.; Ikenaka, T.  
J. Biochem. 79, 1183-1196, 1976  
A:Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino acid sequence of the fragments I and II  
A:Reference number: A91940; MUID:76260153; PMID:956149  
A:Accession: A91940  
A:Molecule type: protein  
A:Residues: 223-288 572-608 <IS2>  
A:Note: 262-Leu was also found  
R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.  
Cancer Res. 38, 3483-3486, 1978  
A:Title: Copper(II)-binding ability of human alpha-fetoprotein.  
A:Reference number: A90758; MUID:79001617; PMID:60265  
A:Contents: annotation; copper binding  
R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.  
J. Immunol. 143, 1680-1684, 1989  
A:Title: Structures of histamine-releasing peptides formed by the action of acid protease  
A:Reference number: A45800; MUID:89341406; PMID:2474609  
A:Accession: C45800  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 166-173 <CAR>  
R:Heard, J.  
Mol. Cell. Biol. 7, 2425-2434, 1987  
A:Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved reporter gene assay  
A:Reference number: I57621; MUID:87286876; PMID:3475566  
A:Accession: I57621  
A:Status: preliminary; translated from GB/EMSL/DBD  
A:Molecule type: DNA  
A:Residues: 1-5 <RES>  
A:Cross-references: GB:M16925; NID:9202828; PIDN:AAA0712.1; PID:9554412  
C:Superfamily: serum albumin; serum albumin repeat homology  
C:Keywords: carrier protein; duplication; metal binding; plasma  
P:1-18/Domain: signal; sequence #status experimental <SIG>  
P:19-24/Domain: propeptide; #status experimental <PRC>  
P:25-608/Product: serum albumin #status experimental <MAT>  
P:29-202/Domain: serum albumin repeat homology <SA1>  
P:221-394/Domain: serum albumin repeat homology <SA2>  
P:413-592/Domain: serum albumin repeat homology <SA3>  
P:27/Binding site: copper (His) #status experimental  
P:77-86, 99-113, 114-125, 148-193, 192-201, 224-270, 269-277, 289-303, 302-313, 340-365, 384-385

Query Match 78.2%; Score 2426; DB 1; Length 608;  
Best Local Similarity 73.4%; Pred. No. 5.6e-153;  
Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;

QY 1 DAHSEVAHRRFKOLGEENPKALVLIAPACYQQCPFDHVKLVNEVTEFAKTCVADSEAE 60  
DB 25 EAHKSEIAHRRNDLGEENFQGLVLIAPSYQYQQCPFDHVKLVNEXTEFAKTCVADSEAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVLRPEV 120  
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDQNPRLVLRPEV 144  
QY 121 DVMTAFHNDNEETPLKKYLYEIAIRRHYPFYAPPELLFFAKYKAAFTSCQAAADKACLLP 180  
DB 145 BAYCTSPQENPTSPFLGHYLHEVARRHPFYAPPELLFFAKYKAAFTSCQAAADKACLLP 204  
QY 181 KDELDRDEGKASSAKORLKCASLQKQFGERAFKAWAVARLSQRFKAEFAEYSKLVTLTK 240  
DB 205 KIDAMREKVLASSAKORLKCASLQKQFGERAFKAWAVARLSQRFKAEFAEYSKLVTLTK 264  
QY 241 VHTCCCHGDLLECCADDRADLAKYICENODSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300  
DB 265 INKCECHGDLLECCADDRADLAKYICENODSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324  
QY 301 DLPSLAADPFVESKDVCKYKAAKQVFLGMFLYEVARRHPDYSVVLRLAKTYETTLTK 360  
DB 325 DLPSLAADPFVESKDVCKYKAAKQVFLGMFLYEVARRHPDYSVVLRLAKTYETTLTK 384  
QY 361 CAADPHCEYKAVDFEKPVLVEEPQNLKQNCLEFEQOLGEYKFNQALLVRYTKKQPOVST 420  
DB 385 CAEDPHCEYKAVDFEKPVLVEEPQNLKQNCLEFEQOLGEYKFNQALLVRYTKKQPOVST 444

Db 262 VHKCCGHDJLLECCADRADLAKYVCENQDTISTKLKCCDCKLLEKSHGCIABAKDELP 321  
 QY 301 DLPSLAADPFVESKDVCKVNAEAKDVFGLMFLYEYARRHPDYSVVLRLRLAKTYETTLK 360  
 Db 322 CLNPLLEHDFVEDNECKVNAEAKDVFGLTFLYESSRRHPDYSVSLLRIRAKVEATUSD 381  
 QY 361 CAADAPHECYAKVDFEFKPLVEEPONLKKONCELFEQLGEYKFNQALLVRYTKKVPQVST 420  
 Db 382 CAKEPPACVATVDFKFPQVJDEPKLKKONCELFEKJGEYQFNALIVRYTKKVPQVST 441  
 QY 421 PTLVEVSRNLKVGSKCKCKHPKAKRKPCCADYLSLVVNLQLCVLHEKTPVSDRVTKCCTES 480  
 Db 442 PTLVEVARKLGVGSRCKKPESELRSCADYLSLVNLRLCVLHEKTPVSEKVTKCTES 501  
 QY 481 LVNRRPFSALVEVDETVVPKEFNAETPTHADICTLSEKEROKKQOTALVELVKKHPKAT 540  
 Db 502 LVNRRPFSALTDDETVYPKEFVEGTFTHADICTLPEDEKQIKKQOTALVELLKKHPAT 561  
 QY 541 KEQLKAVNCDPFAAEVEXCKKADDKETCFAEEGKKLV 576  
 Db 562 EEQRTVLGNPFAAEVQKCCAAPDHEACFAVEGPKFV 597

RESULT 9  
 JC5838  
 albumin - Mongolian Iird  
 C:Species: Meriones unguiculatus (Mongolian Iird)  
 C:Date: 05-Mar-1998 Wsequence\_revision 13-Mar-1998 #text\_change 19-May-2000  
 C:Accession: JC5838  
 R.Yoshida, K.: Seto-Ohshima, A.; Sinohara, H.  
 DNA Res. 4, 351-354, 1997  
 A:Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in  
 A:Reference number: JC5838; MUID:98116663; PMID:9455485  
 A:Accession: JC5838  
 A:Molecule type: mRNA  
 A:Residues: 1-609 <YOS>  
 A:Cross-references: DDBJ:AB006197; NID:G2317277; PIDN:BAA21765.1; PID:G2317278  
 A:Experimental source: liver  
 C:Superfamily: serum albumin; serum albumin repeat homology  
 F:222-395/Domain: serum albumin repeat homology <SA2>

Query Match 76.94; Score 2387; DB 2; Length: 609;  
 Best Local Similarity 73.94; Pred. No. 2,1e-150;  
 Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;

QY 2 AHKSEVHRFDLGEENFKALVLIATFAQYQOCFFEDHVKLVNEVTEFARTCVADESAEN 61  
 Db 27 AHKSEIHRKDLGEYFKGLVLTFSQYLKQSEYEHVKLVREVTFDPSNCAKDESAEN 86  
 QY 62 COKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQHKDQDNPLRLVLRPEVD 121  
 Db 87 COKSLHTLFGDKLCSLPNFGKEYAEMADCCAKQEPERNECFQHKDQDNPLRPFKRAEPD 146  
 QY 122 VMC2AFHONESTFLKYLLEYLARRHPYFVAPELLFTAKRYKAAFTTECCAAAKAAILPK 181  
 Db 147 ANCTAFQENACAFMGVYHEVARHPYFYGPELLYLACKYAVLTTECCAADKGACLTPK 206  
 QY 182 DEURDEGKASAKORLUKASLQKGERAFKAWAVARLSQRFFKAEFASVSKLVLTSLTKV 241  
 Db 209 LPAALKEKALVSAVRQRLKCSMKKFGERAFKAWAVARMSQTFENADPAEITKLATSLTKV 266  
 QY 242 HTECCGDLLECADRADLAKYVCENQDSTSSKLKCCCKPLLEKSHGCIABEVNDMPAD 301  
 Db 267 TOECCGDLLESCADRAELAKYVCENQASTSSKLQACCCKEMIQKSOCLAEVEHDDMPAD 326  
 QY 302 APSLAADPFVESKDVCKVNAEAKDVFGLMFLYEYARRHPDYSVVLRLRLAKTYETTLK 361  
 Db 327 LPALTAQFVEDKQCKVNAEAKDVFGLTFLYESSRRHPDYSVSLLRIRAKVEATLTKK 386  
 QY 362 AAADPHCYAKVDFEFKPLVEEPONLKKONCELFEQLGEYKFNQALLVRYTKKVPQVSTP 421  
 Db 387 AEADPHACYGHVDFEFKPLVEEPONLVKNCLELYKLGEGYQFNALIVRYTKKVPQVSTP 446

```
QY 422 TAJEVSRLGKVGSKCKCHPEAKMPCAEDYLSVYLNQLCVLHEKTPVSORVTKCCTESL 481
DB 447 TAJEARSRLGRVGHCCALPEKXLPVEDYLSALNRPVCLLHEKTPVSQVTKCCSGSL 506
QY 482 VNRPCFSALVETDYTPKFNABTTFTHADICTLSEKERQIKKQALVJLKHKKPKATK 541
DB 507 VERRPCFSALPVDYTPKFNABTTFTHADICTLSEKERQIKKQALVJLKHKKPKATK 566
QY 542 EQLKAVMDGFAAFVEKCKCKADDDKTCFAEBGKGLVAASQAAL 583
DB 567 EQLKAVMDGFAAFVEKCKCKADDDKTCFAEBGKGLVAASQAAL 609
RESULT 10
A05139
serum albumin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: A05139; I48638
R:Minhrett, P.P.; Law, S.W.; Dugaczky, A.
Mol. Biol. Evol. 2, 347-358, 1985
A:Title: The rate of molecular evolution of alpha-fetoprotein: approaches that of pseudogenes
A:Reference number: A93455; MUID:882:6123; PMID:2452956
A:Accession: A05139
A:Molecule type: mRNA
A:Residues: 1-418 <MIN>
A:Cross-references: GB:M16111; NID:G191764; PIDN:AAA37190.1; PID:G191765
R:Boccaccio, C.; Deschatrete, J.; Meunier-Rotival, M.
Gene 88, 181-186, 1990
A:Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the
A:Reference number: I48638; MUID:92269606; PMID:1971802
A:Accession: I48638
A:Status: preliminary; translated from GB/EXBL/DBJ
A:Molecule type: DNA
A:Residues: 379-453 <BOC>
A:Cross-references: EMBL:X13060; NID:G52939; PIDN:CAA1458.1; PID:8899334
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-104/Domain: serum albumin repeat homology (fragment) <SA1>
F:123-296/Domain: serum albumin repeat homology <SA2>
F:315-453/Domain: serum albumin repeat homology (fragment) <SA3>
Query Match 60.0%; Score 1861; DB 2; Length 453;
Best Local Similarity 72.2%; Pred. No. 9.5e-116;
Matches 327; Conservative 64; Mismatches 62; Indels 0; Gaps 0;
QY 75 CTVALTRETGEMADCCAKQEPERNECFLOHKQDNPQLP:VRPEVVMCTAFHNEET 134
DB 1 CAIENLRENYGEJADCCCTKQEPERNECFLOHKQDNPQLP:VRPEVVMCTAFHNEET 134
QY 135 LKKYLYEIAIRHPHYFYAPELLFFAKRYKAAFTGCCQAADKAAACLLPKLDELDEGRASSA 194
DB 61 MGHVLYEVARHPHYFYAPELLFYAYAOYNE-LTCCAEADKESCLTPKLDGVKEALVSSV 130
QY 195 KQPLKASLQKFGREAFKAWAVARLSQRPKAEPAEVSXVTDUTKYVHTCCGGDLLECA 254
DB 121 RQRMKCSMKGFERAFKAWAVARLSQTFNADPAE:TKLATLTKVKECCGGDLLECA 180
QY 255 DDADLAKYCENQDSISSKLEKCEKPLLEKSKCIAEVNDENPADLP:AAQFVESKD 314
DB 181 DDRAELAKYCENQATISSKLTQCCDPLKKAHCLSEVHD:TPALPAIADFVEDQE 240
QY 315 VCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKCCAAADPHCEYAKVF 374
DB 241 VCKNYAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKCCAAENPACVGTVL 300
QY 375 DEFKPLVEEPQNLKQCELFQELGEYKFNQNALIVRYTKVPQVSTPTLVSEVRNLKGVG 434
DB 301 AEFQPLVEEPKLVKTNCDYKELGEYGFQNALIVRYTKVPQVSTPTLVSEVRNLKGVG 360
QY 435 SKCKHPEAKMPCAEDYLSVYLNQLCVLHEKTPVSORVTKCCTESLWNRPCFSALV 494
DB 361 TKCCTLPEDQRLPCVEDYLSALNRPVCLLHEKTPVSQVTKCCTESLWNRPCFSALV 420
```

```
QY 495 ETVVPKEAETETPHADICTLSEKERQIKKOT 527
DB 421 ETVVPKEAETETPHADICTLSEKERQIKKOT 453
RESULT 11
ABCHS
serum albumin precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S15571; A05078; A13451
R:Cassady, A.I.; Salkild, C.K.; Baverstock, P.; Wallace, J.C.
submitted to the EMBL Data Library, July 1991
A:Reference number: S15571
A:Accession: S15571
A:Molecule type: mRNA
A:Residues: 1-615 <CAS>
A:Cross-references: EMBL:X60688; NID:G63747; PIDN:CAA43098.1; PID:G63748
R:Hache, R.J.G.; Wiskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.
J. Biol. Chem. 258, 4556-4564, 1983
A:Title: The 5' noncoding and flanking regions of the avian very low density apolipoprotein
A:Reference number: A05078; MUID:83161037; PMID:6187737
A:Accession: A05078
A:Molecule type: DNA
A:Residues: 1-28 <HAC>
A:Cross-references: GB:V00381; NID:G63038; PIDN:CAA23680.1; PID:G63039
R:Rozer, A.M.; Geller, D.M.
Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977
A:Title: Chicken microsome albumin: amino terminal sequence of chicken proalbumin
A:Reference number: A13451; MUID:78019943; PMID:911327
A:Accession: A13451
A:Molecule type: protein
A:Residues: 19-23, 25-30 <ROS>
A:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copious
monomers (weak bonds) with these hormones promote their transfer across the membranes.
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-26/Domain: propeptide #status predicted <PRO>
F:27-613/Product: serum albumin #status predicted <MAT>
F:32-206/Domain: serum albumin repeat homology <SA1>
F:225-398/Domain: serum albumin repeat homology <SA2>
F:417-596/Domain: serum albumin repeat homology <SA3>
F:3/Binding site: cooper (H:S) #status predicted
F:80-89, 102-118, 117-128, 152-197, 196-205, 228-274, 273-281, 293-307, 306-317, 344-389, 388-
Query Match 50.2%; Score 1557.5; DB 1; Length 615;
Best Local Similarity 46.7%; Pred. No. 1.6e-95;
Matches 273; Conservative 118; Mismatches 192; Indels 1; Gaps 1;
QY 3 HKSEVAHRPKDLGEGNFKALV:JFAQYLQCCPFEDHVKLVNEVTEFAKTCVADSAENC 62
DB 3C HKSEIAHRYNDUKETTFKAVAM:TFQYLCRCYEGLSKLVKDVYDLSAQKCVANSDAPEC 89
QY 63 DKSLATLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKQDNPQLP:LRPEVD 121
DB 90 SKPLPSIILDEICQVEKURDYSYGAMADCCSKADPERNECFLSFKVSPQDFVQYQRPASD 149
QY 122 VMCTAFPHNEETFLKKYLYEIAIRHPHYFYAPELLFFAKRYKAAFTGCCQAADKAAACLLPK 181
DB 150 VICQYQDNRVSFGLHFYISVARRHPFLYAPAILSFVDFEHALQSCCKESVGGACLDTK 209
QY 182 LDELDEGRASSAKQRLKAS:OKGGERAFKAWAVARLSQRPKAEPAEVSXVTDUTKYV 241
DB 210 EVMREKAKGVSVKQYFCGILKQFGRVQFARQYIYLSQKYPKAPFSEVSXVHDS:GV 269
QY 242 HTBCCGCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSKSHCIAEVNDENPAD 301
DB 270 HKECEGDMVCEMDNARMNMLCSQDDVFSOKIKDCEKPIVERSQCIIMEAFPEKPAD 329
QY 302 LPSLAADPVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKCC 361
DB 361 LPSLAADPVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKCC 361
```

```

Db 330 LPSLVKZY:EDKEVCKSPFAGHDAFMAEFVVEYSRRHPEFSIQL-WRIAKGVESLLEKCC 389
Qy 362 AAADPHCYAKVDFDEKPLVVEBPQN::KNCBSEFQOLGEYKFNQALLVRYTKVPQVSTP 421
Db 390 KTDNPAECYANAQEQUNQHIKETQDVVTKNCDLLDHHGEADFEKGI-LRYTKVPQVSTP 449
Qy 422 TLVEVSRLNGKVGSKCKPEAKRMPCAEDY-SVVLNQLCVLHEKTPVSDRYVCKCTESL 481
Db 450 LLETKGKMTTIGKCCQGGEDRMACSEGYLSIVHDTCRQETTPNINNVQCCSGLY 509
Qy 482 VNRRPCFSALEVDYVYKFEKNAETTFHAD-CTLSEKEROIKKQTA-VLVLKHKPKATK 541
Db 510 ANRRPCFTANGVDTKVVPFPMCPMESPEKLCAPAEEREVGQMKLLNLNJKRQPMTE 569
Qy 542 EQKAVMDQFAFVVEKCKKADQKTCFAESGKKLVAAASOAAALGL 585
Db 570 EQIKTIADQFTAMVDKCKQSQDINTCFEGEGANLVQSSRATLGI 633

RESULT 12
JC4258
alpha-fetoprotein precursor - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 27-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 20-Aug-1999
C:Accession: JC4258
R:Nishio, H.; Gibbs, P.E.M.; Xinghetti, P.P.; Zielinski, R.; Dugaiczky, A.
Gene 182, 213-220, 1995
A>Title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to
A:Reference number: JC4258; MUID:96012345; PMID:7557431
A:Accession: JC4258
A:Molecule type: DNA
A:Residues: 1-609 <NIS>
A:Cross-references: GS:U2106; NID:G84131; PID:AAA1641; PID:8841312
C:Comment: This protein is a plasma protein produced in the fetal and neonatal liver and
c similar properties and structure.
C:Genetics:
A:Gene: afp
A:Map position: 3p
A:Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551/3
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: glycoprotein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-609/Product: alpha-fetoprotein #status predicted <MAT>
F:26-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:42,251/Binding site: carbohydrate (Asn) [covalent] #status predicted

Query Match 40.4%; Score 1253.5; DB 2; Length 609;
Best Local Similarity 40.4%; Pred. No. 2e-75;
Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;

Qy 3 HKSE-----VAHFKDLGSENFALVLIAPQV:QCCPDHVKLVNVEFAKTCVADE 57
Db 22 HRNEYGIASLDYQCTAEINLTDJAT:FFAQVQVQATYKEVSKYKQALTA:EKPTGDE 8:
Qy 58 SAENCDSJLT*LFQCKACTVATIRETYGEMADCCAKQEPERNECFQHKDCKP-NLPLRV 116
Db 82 QSAGCLNQVPAFLEE-CREKEILKRVGH-SDCCSQSEGRHNCFLAHKKPTASIPFQ 140
Qy 117 RPEVDVVTAFHDNEFTLKKYLYEARRHPYVAPPELLFAKRYKAAATECCQAADKAA 176
Db 141 VPEPVTSCAEYEDRETFMKNF:YEIARRHPFLYAPT:ILWAARYSKIPSCSKAENAVE 200
Qy 177 CLLPKLDELDEGKASAKORLKAS:CKFGERAFKAWAVAR:SQPFFKAEFNEVSKLVT 236
Db 201 CFQTKAATVTVKELRESSLNQHACAVNKNFSTRTFQ:ATVTKLSQRF:KVN:FEIQKLY 260
Qy 237 DLTQVH:ECCHGDLLECADRAD:AKYICGNQDSISKLECC:EXPLLEKSHKIAEVEN 296
Db 261 DVAVHVEHCCGSDVLDQJGKIMSYICQQDTLSNKTTECCKLTTLLEGQCIIHAEND 320
Qy 297 EMPADLPSLAADFVESKDVCKNVAEAKDVFGLMFLVEYARRHPYVSVVLLRLAKTYET 356

```

```

Db 321 EKPEG:SPNLRFLGDRDNQFSSGEKNIFLASFVHEYSRRHPQLAVSVILRAKGYQEL 380
Qy 357 LEKCCAAAPHECYAKVDFDEKPLVVEBPQNLIKONCELEFEQGEYKFNQALLVRYTKVP 416
Db 381 LEKCFOTENPLEQDQGBEELQKYIQESQALAKRSQGLFQKLGVEYVLQNAF:VAVTKKAP 440
Qy 417 QVSTPTLVEVSRLNGKVGSKCKKHPEAKRMPCAEDY-SVVLNQLCVLHEKTPVSDRYVCK 476
Db 441 QJSSSEMAITRQMAATAATCCOLSEDKLLACGEGAAI:IGHLCIRHETTPVNPQVQCC 500
Qy 477 CTESLVNRRPCFSALEVDYVYKFEKNAETTFHADICTLSEKEROIKKQTA-VLVLKHK 536
Db 501 CTSSYANRRPCFSALLVDETYVPPAFSDSKF:FRKDLCOAGVALQTMKQBEF:INLVKQK 560
Qy 537 PKATKQLKAVMDQFAFVVEKCKKADQKTCFAESGKKLVAAASOAAALGL 585
Db 561 PQ:TESQLEAV:ADFSGLLEKCKQCGEQEVCFAESGQK:SKTRAAALGV 609

RESULT 13
FPHU
alpha-fetoprotein precursor [validated] - human
N:Alternate names: AFP; alpha-1-fetoprotein; alpha-fetoglobulin
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 08-Dec-2000
C:Accession: A26624; S37655; A93961; A91497; A23699; A61480; A90624; A90757; A93042;
R:Gibbs, P.E.M.; Zielinski, R.; Boyd, C.; Dugaiczky, A.
Biochemistry 26, 1332-1343, 1987
A>Title: Structure, polymorphism, and novel repeated DNA elements revealed by a comp
A:Reference number: A26624; MUID:87185438; PMID:2436661
A:Accession: A26624
A:Molecule type: DNA
A:Residues: 1-609 <GIB>
A:Cross-references: GS:M16110; NID:G773678; PIDN:AAB38754.1; PID:gl79236
R:McVey, J.H.; Michaelides, K.; Hansen, L.P.; Ferguson-Smith, M.; Tilghman, S.; Krum
Hum. Mol. Genet. 2, 379-384, 1993
A>Title: A G-A substitution in an HNF I binding site in the human alpha-fetoprotein
A:Reference number: S37655; MUID:93278385; PMID:7684942
A:Accession: S37655
A:Molecule type: DNA
A:Residues: 1-28 <MCV>
A:Cross-references: EMBL:Z19532; NID:G28527; PIDN:CAA79592.1; PID:G28528
A>Note: The authors translated the codon TAT for residue 26 as Thr
R:Morinaga, T.; Sakai, M.; Wegmann, T.G.; Tamaoki, T.
Proc. Natl. Acad. Sci. U.S.A. 80, 4604-4608, 1983
A>Title: Primary structures of human alpha-fetoprotein and its mRNA.
A:Reference number: A93961; MUID:83273664; PMID:6192439
A:Accession: A93961
A:Molecule type: mRNA
A:Residues: 1-609 <MOR>
A:Cross-references: GS:J00077; NID:G311348; PIDN:CAA24758.1; PID:G31351
R:Beattie, W.S.; Dugaiczky, A.
Gene 20, 445-422, 1982
A>Title: Structure and evolution of human alpha-fetoprotein deduced from partial seq
A:Reference number: A91497; MUID:83158778; PMID:6187626
A:Accession: A91497
A:Molecule type: mRNA
A:Residues: 429-556 <BEA>
A:Cross-references: GS:J00076
R:Fucini, P.; Siciliano, R.; Malorni, A.; Marino, G.; Tecce, M.F.; Ceccarini, C.; Ter
Biochemistry 30, 5061-5066, 1991
A>Title: Human alpha-fetoprotein primary structure: a mass spectrometric study.
A:Reference number: A23699; MUID:91242409; PMID:1709810
A:Accession: A23699
A:Molecule type: protein
A:Residues: 19-45:60-97:102-107:122-184:187-249:255-489:507-609 <FUC>
R:Tecce, M.F.; Terrana, B.; Giuliani, M.M.; Ceccarini, C.
J. Nucl. Med. Allied Sci. 34, 213-216, 1990
A>Title: Characterization of in vitro expressed human alpha-fetoprotein as highly re
A:Reference number: A61480; MUID:91225826; PMID:1709209
A:Accession: A61480
A:Molecule type: protein

```

Query Match:	40.3%	Score	1249.5;	DB 1;	Length	609;		
Best Local Similarity	39.9%;	Pred.	No. 3.8e-75;					
Matches	235;	Conservative	116;	Mismatches	231;	Indels	7; Gaps	3;
Qy	3	HKSE-----VAHRRKDLGEENFKALVLIAPAYQLQCQPFEDRVKLVNNEVTFAKTCTVADE	57					
Db	22	HRNEYGIAS::LDSYQCTAEISLAIDLAT::FFAQFYQEATYKEVSQWAXDALTEAKTPGDE	81					
Qy	58	SAENCDSK::HTLPDGKLCTVATLRETGYEMADCCAKOEPERNECFLOHKDNP-NLFRLV	116					
Db	82	QSGGLENQLPAFLLEELCHEXEILEKYGH-SDCSSQSEGRHNCFLAHKKPTPASIPLFQ	143					
Qy	117	RPEVVVMCTAHDNNEETFLKYLVEIARRHPDYFAPELLFNAKRYKAAFTECCQAADKAA	176					

	Query Match:	40.08;	Score	1242.5;	DB	1;	Length	609;
	Best Local Similarity	39.64;	Pred.	No. 1.1e-74;				
	Matches	233;	Conservative	117;	Mismatches	232;	Indels	7; Gaps
Qy	3	HKSE----	VAHRPKDLGEGNFKALV--	AQAQYLQCQFFEDHVKLVNNEVFTEFAKTCVADE	57			
Dd	22	HRNEYGIASLDSYQCAEISLADLTIFAQFYQEATYKEVSNMVKDALTAEIKPTGDE	81					
Qy	58	SAENCDSKLHTLFQDKLCITVATRETGVENADCACAKOEPERNECFLOHKDDNP-NLPRLV	116					
Dd	82	QSACCLNQIPAFLEULCHEKEIILEKYG-LSDCCSQSEGRHN-CFLAHKKPTPASIPUFG	140					
Qy	117	RPEVDVNCTAFPHDNNEETFLKKLYIEIARRHFYFAPPELLFPAYKAAFTTECCQAADKAA	176					

Search completed: October 27, 2003, 15:21:55  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 15:08:38 ; Search time 25 Seconds  
(without alignments)  
1100.425 Million cell updates/sec

Title: US-09-833-117-18

Perfect score: 3103

Sequence: 1 DAHKEVAHRFKDLGEENFK.....TCFAEEGKKLVASQAALGL 565

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	3103	100.0	609	1	ALBU_HUMAN	P02768	homo sapien
2	2342	94.8	600	1	ALBU_MAMMU	Q28522	macaca mulia
3	2320	84.4	608	1	ALBU_FELCA	P49064	felis silve
4	2562	82.6	608	1	ALBU_CANFA	P49022	canis fami
5	2475.5	79.8	607	1	ALBU_HORSE	P35747	equus cabai
6	2450.5	79.0	607	1	ALBU_BOVIN	P02769	bos taurus
7	2446	78.8	608	1	ALBU_RABIT	P49065	oryctolagus
8	2432.5	78.4	607	1	ALBU_SHEEP	P14639	cvis aries
9	2426	78.2	608	1	ALBU_RAT	P02770	rattus norv
10	2411.5	77.7	605	1	ALBU_PIG	P08835	sus scrofa
11	2397	76.9	609	1	ALBU_MERUN	O35090	meriones un
12	2378	76.6	609	1	ALBU_MOUSE	P07724	mus musculu
13	1557.5	50.2	615	1	ALBU_CHICK	P19121	gallus gall
14	1253.5	40.4	609	1	FETA_PANTR	Q28769	pan treglod
15	1249.5	40.3	609	1	FETA_HUMAN	P02771	homo sapien
16	1242.5	40.0	609	1	FETA_GORGO	P28050	gorilla gor
17	1205	38.8	607	1	ALB2_XENLA	P14872	xenopus lae
18	1200	38.7	609	1	FETA_HORSE	P43066	equus cabai
19	1164.5	37.5	606	1	ALB1_XENLA	P08759	xenopus lae
20	1084	34.9	605	1	FETA_MOUSE	P02772	mus musculu
21	1067	34.4	611	1	FETA_RAT	P02773	rattus norv
22	1055	34.0	599	1	AFAM_HUMAN	P43652	homo sapien
23	944	30.4	611	1	AFAM_MOUSE	O83020	mus musculu
24	928	29.9	608	1	AFAM_RAT	P36953	rattus norv
25	747.5	24.1	608	1	ALB1_SALSA	P21848	salmo salar
26	742.5	23.9	608	1	ALB2_SALSA	Q03156	salmo salar
27	699	22.5	182	1	ALBU_RANCA	P21847	rana catesb
28	440.5	14.2	1423	1	ALBU_PETMA	Q90274	petronysen
29	386	12.4	474	1	VTDB_HUMAN	P02774	homo sapien
30	381	12.3	476	1	VTDB_RAT	P04276	rattus norv
31	378	12.2	476	1	VTDB_RABIT	P53789	oryctolagus
32	372	12.0	472	1	VTDB_MOUSE	P21614	mus musculu
33	151.5	4.9	1605	1	RRB1_MOUSE	Q99p15	mus musculu

34	144.5	4.7	8797	1	SNEL_HUMAN	Q8nf91	homo sapien
35	138.5	4.5	1410	1	RRB1_HUMAN	Q9p289	homo sapien
36	133.5	4.3	1391	1	NST3_DROHY	Q08696	drosophila
37	132.5	4.3	2230	1	GOG4_HUMAN	Q13439	homo sapien
38	129.5	4.2	1972	1	MYHB_HUMAN	P15749	homo sapien
39	129	4.2	3210	1	CENF_HUMAN	P49454	homo sapien
40	128	4.1	1005	1	RA50_METJA	Q58718	methanococc
41	126.5	4.1	1972	1	MYHB_RABIT	P35748	oryctolagus
42	126	4.1	1189	1	SMC2_CHICK	Q90988	gallus gall
43	126	4.1	3259	1	GIAN_HUMAN	Q14789	homo sapien
44	125	4.0	3038	1	TRIO_HUMAN	O75962	homo sapien
45	124.5	4.0	1790	1	USO1_YEAST	P25386	saccharomyc

## ALIGNMENTS

RESULT 1	
ALBU_HUMAN	
STANDARD:	PRC: 609 AA.
P02768; Q95574; Q13140; Q9P157; Q9P117; Q9H8S3; Q9JUZC;	
21-JUL-1986 (Rel. 51, Created)	
01-APR-1990 (Rel. 14, Last sequence update)	
15-SEP-2003 (Rel. 42, Last annotation update)	
Serum albumin precursor.	
ALB.	
Homo sapiens (Human).	
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
NCBI_TaxID=9606;	
[1]	
SEQUENCE FROM N.A.	
MEDLINE=86296112; PubMed=3009475;	
Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,	
Beattie W.G., Dugaiczky A.;	
"Molecular structure of the human albumin gene is revealed by	
nucleotide sequence within g11-22 of chromosome 4.";	
J. Bio. Chem. 261:6747-6757(1986);	
[2]	
SEQUENCE FROM N.A., AND VARIANT LYS-420.	
MEDLINE=82081882; PubMed=6171778;	
Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,	
Najarian R.C., Seeburg P.H., Wion K.L.;	
"The sequence of human serum albumin cDNA and its expression in E.	
coli.";	
Nucleic Acids Res. 9:6103-6114(1981);	
[3]	
SEQUENCE FROM N.A., AND VARIANT GLY-121.	
MEDLINE=82105994; PubMed=6275391;	
Dugaiczky A., Law S.W., Dennison O.E.;	
"Nucleotide sequence and the encoded amino acids of human serum	
albumin mRNA.";	
Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982);	
[4]	
SEQUENCE FROM N.A.	
TISSUE=Liver;	
Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;	
Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.	
[5]	
SEQUENCE FROM N.A. (PROCC903/PRO1708/PRO2C44/PRO2619/PRO2675).	
TISSUE=Fetal liver;	
Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,	
Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;	
"Functional prediction of the coding sequences of 121 new genes	
deduced by analysis of cDNA clones from human fetal liver.";	
Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.	
[6]	
SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.	
Huang M.C., Wu H.T.;	
"The cDNA sequences of human serum albumin.";	
Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.	
[7]	
SEQUENCE FROM N.A.	

TISSUE=liver, and Skeletal muscle;  
 MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge C.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Ziegler B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong J.,  
 RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uslan T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting X., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Snailus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [8]  
 RP SEQUENCE OF 25-609.  
 RP MEDLINE=76187907; PubMed=1225573;  
 RA Meloun B., Moravsek L., Kostka V.;  
 RT "Complete amino acid sequence of human serum albumin.";  
 RL FEBS Lett. 58:134-137(1975).  
 [9]  
 RP SEQUENCE OF 25-609.  
 RP Brown C.R., Shockley P., Behrens P.Q.;  
 RL [in:] Bing D.H. (eds.);  
 RT "The chemistry and physiology of the human plasma proteins, pp.23-40,  
 RL Pergamon Press, New York (1979).  
 [10]  
 RP SEQUENCE OF 1-455 FROM N.A.  
 RP TISSUE=Liver;  
 RA Menaya J., Parrilla R., Ayuso X.S.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 [11]  
 RP SEQUENCE OF 1-26 FROM N.A.  
 RP MEDLINE=8640099; PubMed=2419329;  
 RA Urano Y., Watanabe K., Sakai X., Tamaoki T.;  
 RT "The human albumin gene. Characterization of the 5' and 3' flanking  
 RT regions and the polymorphic gene transcripts.";  
 RL J. Biol. Chem. 261:3244-3251(1986).  
 [12]  
 RP SEQUENCE OF 222-229.  
 RP MEDLINE=76257808; PubMed=955075;  
 RA Walker J.E.;  
 RT "Lysine residue 199 of human serum albumin is modified by  
 RT acetylsalicylic acid.";  
 RL FEBS Lett. 66:173-175(1976).  
 [13]  
 RP SEQUENCE OF 25-44 AND 480-499.  
 RP TISSUE=Heart;  
 RA MEDLINE=95203287; PubMed=7895732;  
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
 RT "The human myocardial two-dimensional gel protein database: update  
 RT 1994.";  
 RL Electrophoresis 15:1459-1465(1994).  
 [14]  
 RP DISULFIDE BONDS.  
 RA Saber N.A., Stockbauer P., Moravsek L., Meloun B.;  
 RT "Disulfide bonds in human serum albumin.";  
 RL Collect. Czech. Chem. Commun. 42:564-579(1977).  
 [15]  
 RP BILIRUBIN-BINDING SITE.  
 RP MEDLINE=78186630; PubMed=656055;  
 RA Jacobsen C.;  
 RT "Lysine residue 240 of human serum albumin is involved in high-  
 RT affinity binding of bilirubin.";  
 RL Biochem. J. 171:453-459(1978).  
 [16]  
 RP VARIANT CANTERBURY ASN-337.  
 RP MEDLINE=87157744; PubMed=3828358;  
 RA Brennan S.O., Herbert P.;  
 RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second  
 RT domain of serum albumin.";  
 RL Biochim. Biophys. Acta 912:191-197(1987).  
 [17]  
 RP VARIANTS NAG-2 AND NAG-3.  
 RP MEDLINE=8068523; PubMed=3479777;  
 RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,  
 RA Satoh C., Neel J.V.;  
 RT "Amino acid substitutions in inherited albumin variants from  
 RT Amerindian and Japanese populations.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8031-8005(1987).  
 [18]  
 RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.  
 RP MEDLINE=89345611; PubMed=2762316;  
 RA Arai K., Madison J., Huss K., Ishioke N., Satch C., Fujita M.,  
 RA Neel J.V., Sakurabayashi I., Putnam F.W.;  
 RT "Point substitutions in Japanese albumins.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).  
 [19]  
 RP VARIANTS KANAU; OSAGA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.  
 RP MEDLINE=90115905; PubMed=2404284;  
 RA Arai K., Madison J., Shimizu A., Putnam F.W.;  
 RT "Point substitutions in albumin genetic variants from Asia.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).  
 [20]  
 RP DESCRIPTION OF VARIANT REDHILL.  
 RP MEDLINE=90115852; PubMed=2104980;  
 RA Brennan S.O., Wyles T., Peach R.J., Donaldson D., George P.M.;  
 RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of  
 RT human serum albumin whose precursor has an aberrant signal peptidase  
 RT cleavage site.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).  
 [21]  
 RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBC VALENTIA LYS-106.  
 RP MEDLINE=91062352; PubMed=2247440;  
 RA Galliano M., Mirchiotti L., Porta F., Rossi A., Ferri G., Madison J.,  
 RA Watkins S., Putnam F.W.;  
 RT "Mutations in genetic variants of human serum albumin found in  
 RT Italy.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).  
 [22]  
 RP VARIANT VENEZIA.  
 RP MEDLINE=91296740; PubMed=2068071;  
 RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,  
 RA Mirchiotti L., Putnam F.W.;  
 RT "A donor splice mutation and a single-base deletion produce two  
 RT carboxyl-terminal variants of human serum albumin.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:59:9-5963(1991).  
 [23]  
 RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;  
 RP KOMAGOME-2 ARG-152 AND KOMAGOME-1 GIU-396.  
 RP MEDLINE=92052189; PubMed=1946412;  
 RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,  
 RA Matsuda Y.-i., Amaki I., Putnam F.W.;  
 RT "Genetic variants of serum albumin in Americans and Japanese.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).  
 [24]  
 RP VARIANT CASEBROOK ASN-518.  
 RP MEDLINE=93316157; PubMed=1859851;  
 RA Peach R.O., Brennan S.O.;  
 RT "Structural characterization of a glycoprotein variant of human serum  
 RT albumin: albumin Casebrook (494 Asp-->Asn)."  
 RL Biochim. Biophys. Acta 1097:49-54(1991).  
 [25]  
 RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.  
 RP MEDLINE=92190239; PubMed=1347703;  
 RA Mirchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,  
 RA Rochu D., Porta F.;  
 RT "Two albumins with identical electrophoretic mobility are produced  
 RT in the same liver."

Query Match 100.0%; Score 3103; DB 1; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-184;  
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDLGSENFKALVLIATAFYQJQCPEFHVKLVNVEVTEFAKTCVADESSE 60  
 DB 25 DAHKSVAHRFKDLGSENFKALVLIATAFYQJQCPEFHVKLVNVEVTEFAKTCVADESSE 84  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPVLRVREPV 120  
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPVLRVREPV 144  
 QY 121 DVNCTAFPHNEETFLKKYLVEYARRHPYFAPPELLFFAKRYKAATFECQAAKACALLP 180  
 DB 145 DVNCTAFPHNEETFLKKYLVEYARRHPYFAPPELLFFAKRYKAATFECQAAKACALLP 204  
 QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQFPKAEFAEVSCLVDTLTK 240  
 DB 205 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQFPKAEFAEVSCLVDTLTK 264  
 QY 241 VHTTECHGDLLECCADRADLAKYICENQDISISKLKECCCKPILLESKSHCIAEVENDEMPA 300  
 DB 265 VHTTECHGDLLECCADRADLAKYICENQDISISKLKECCCKPILLESKSHCIAEVENDEMPA 324  
 QY 301 DLPSLAADFVSKDCVKCYAEAKDVFLGMFLVEYARRHPDYVSVLLRLAKTYETLTKC 360  
 DB 325 DLPSLAADFVSKDCVKCYAEAKDVFLGMFLVEYARRHPDYVSVLLRLAKTYETLTKC 384  
 QY 361 CAADAPHECYAKVDFEFLPLVEEPQNLKQNCLEFQGLGEYKFNQALLVRYTKVPQVST 420  
 DB 385 CAADAPHECYAKVDFEFLPLVEEPQNLKQNCLEFQGLGEYKFNQALLVRYTKVPQVST 444  
 QY 421 PTLVEVSRNLGKYGSKCKRHPKAKRMPCAEDYLSVVLNGLCVLHEKTPVSDRYTKCCTES 480  
 DB 445 PTLVEVSRNLGKYGSKCKRHPKAKRMPCAEDYLSVVLNGLCVLHEKTPVSDRYTKCCTES 504  
 QY 481 LVNRPCFSALEVDYTVPKFEKNAETFTPHADICTLSEKERIKKOTALVELVHKPKAT 540  
 DB 505 LVNRPCFSALEVDYTVPKFEKNAETFTPHADICTLSEKERIKKOTALVELVHKPKAT 564  
 QY 541 KEOLKAWMDPFAAFVKKCCADDKETCFABEGKKLVAAQAALGL 585  
 DB 565 KEOLKAWMDPFAAFVKKCCADDKETCFABEGKKLVAAQAALGL 609

## RESULT 2

ALBU\_MACMU STANDARD; PRF: 600 AA.  
 AC Q28522;  
 AT 01-NOV-1997 (Rel. 35, Created);  
 DT 01-NOV-1997 (Rel. 35, Last sequence update);  
 DT 28-FEB-2003 (Rel. 41, Last annotation update);  
 DE Serum albumin precursor (Fragment);  
 GN A.B.  
 OS Macaca mulatta (Rhesus macaque);  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 CC Cercopithecinae; Macaca.  
 CX NCBI\_TaxID:9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93211971; PubMed=8460152;  
 RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,  
 RA Dwalet J., Putnam F.W.;  
 RC "cDNA and protein sequence of polymorphic macaque albumins that differ  
 in bilirubin binding";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).  
 CC -!- FUNCTION: Serum albumin; the main protein of plasma, has a good  
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: V90463; AAA36906.1; --  
 DR PIR: A47391; A47391.  
 DR HSP: P02769; IE78.  
 DR InterPro: IPR000264; Serum albumin.  
 DR Pfam: PF00273; transport prot.; 3.  
 DR PRINTS: PR00802; SERUMALBUMIN.  
 DR ProDom: PDC02486; Serum albumin; 1.  
 DR SMART: SM00103; ALBUMIN\_3.  
 DR SMART: PS00212; ALBUMIN\_3.  
 DR PROSITE: PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 FT SIGNAL 1 10 BY SIMILARITY.  
 FT PROPEP 11 16 BY SIMILARITY.  
 FT CHAIN 17 600 SERUM ALBUMIN.  
 FT DOMAIN 17 197 ALBUMIN 1.  
 FT DOMAIN 204 389 ALBUMIN 2.  
 FT DOMAIN 396 587 ALBUMIN 3.  
 FT BINDING 256 256 COPPER (BY SIMILARITY);  
 FT BILIRUBIN 69 78 BILIRUBIN (POTENTIAL);  
 FT DISULFID 91 107 BY SIMILARITY.  
 FT DISULFID 106 117 BY SIMILARITY.  
 FT DISULFID 140 185 BY SIMILARITY.  
 FT DISULFID 194 193 BY SIMILARITY.  
 FT DISULFID 216 262 BY SIMILARITY.  
 FT DISULFID 261 269 BY SIMILARITY.  
 FT DISULFID 281 295 BY SIMILARITY.  
 FT DISULFID 294 305 BY SIMILARITY.  
 FT DISULFID 332 377 BY SIMILARITY.  
 FT DISULFID 376 385 BY SIMILARITY.  
 FT DISULFID 408 454 BY SIMILARITY.  
 FT DISULFID 453 464 BY SIMILARITY.  
 FT DISULFID 477 493 BY SIMILARITY.  
 FT DISULFID 492 503 BY SIMILARITY.  
 FT DISULFID 530 575 BY SIMILARITY.  
 FT DISULFID 574 583 BY SIMILARITY.  
 SQ SEQUENCE 600 AA; 67880 MW; 545C871A670E740B CRC64;

Query Match 94.8%; Score 2942; DB 1; Length 600;  
 Best Local Similarity 93.5%; Pred. No. 4.8e-184;  
 Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDLGSENFKALVLIATAFYQJQCPEFHVKLVNVEVTEFAKTCVADESSE 60  
 DB 17 DTHKSEVAHRFKDLGSENFKALVLIATAFYQJQCPEFHVKLVNVEVTEFAKTCVADESSE 76  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPVLRVREPV 120  
 DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNPVLRVREPV 136  
 QY 121 DVNCTAFPHNEETFLKKYLVEYARRHPYFAPPELLFFAKRYKAATFECQAAKACALLP 180  
 DB 137 DVNCTAFPHNEETFLKKYLVEYARRHPYFAPPELLFFAKRYKAATFECQAAKACALLP 196  
 QY 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQFPKAEFAEVSCLVDTLTK 240  
 DB 197 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQFPKAEFAEVSCLVDTLTK 256  
 QY 241 VHTTECHGDLLECCADRADLAKYICENQDISISKLKECCCKPILLESKSHCIAEVENDEMPA 300  
 DB 257 VHTTECHGDLLECCADRADLAKYICENQDISISKLKECCCKPILLESKSHCIAEVENDEMPA 316

QY 301 DLPSSAADPVESKGVCKNYAEKXDFVLFQMFLEYEARHPDYSVYLLRLAKTYETILEK 360  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 317 DLPSSAADPVESKGVCKNYAEKXDFVLFQMFLEYEARHPDYSVYLLRLAKAYEATLEK 376  
 QY 361 CAAADPHECVAKVDFBPKLVEBPONLIVONCELFQOLGEYKFNALLVRYTKVPOVST 420  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 377 CAAADPHECVAKVDFBPKLVEBPONLIVONCELFQOLGEYKFNALLVRYTKVPOVST 436  
 QY 421 PTLVEVSRNLGKVGSKCKHPKAKSCAEDYLSVYVNLGLCVHEKTPVSDRVTKCCTES 480  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 437 PTLVEVSRNLGKVGSKCKHPKAKSCAEDYLSVYVNLGLCVHEKTPVSEKVTKCTES 496  
 QY 481 LVNRRPCFSALVDEYVVPKFNATFTTFAADCTILSEKERQIKKOTATLVEVVKHKPKAT 540  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 497 LVNRRPCFSALVDEYVVPKFNATFTTFAADCTILSEKERQIKKOTATLVEVVKHKPKAT 556  
 QY 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAAASQAL 583  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 557 KEQLKGVMDNFAAFVEKCKKADDKETCFABEGPKFVAASQAL 599

## RESULT 3

ALBU\_FELCA STANDARD; PRT; 608 AA.

AC P49064;  
 DT 01-FEB-1996 (Rel. 33, Created;  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor (Allergen Fel d 2).  
 GN ALB.  
 OS Felis silvestris catus (Cat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 OX NCBI\_TaxID=9685;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96194824; PubMed=8647469;  
 RA Hilger C., Grigioni F., Köhnen M., Hentges F.;  
 RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";  
 RL Gene 169,295-296(1996).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Plasma.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X84842; CAA59279.1; -  
 DR PIR; J04560; S57632.  
 DR HSP; P02768; IZ57.  
 DR InterPro; IPR000264; Serum\_albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR PRINTS; PR00802; SERJNALBUMIN.  
 DR ProDom; PD002486; Serum\_albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 DR Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.  
 KW SIGNAL 1 18  
 FT PROPEP 19 24 BY SIMILARITY.  
 FT CHAIN 25 608 SERUM ALBUMIN.  
 FT DOMAIN 25 205 ALBUMIN 1.  
 FT 212 397 ALBUMIN 2.

FT DOMAIN 404 595 ALBUMIN 3.  
 FT METAL 27 27 COPPER.  
 FT DISULFID 77 86 BY SIMILARITY.  
 FT DISULFID 99 115 BY SIMILARITY.  
 FT DISULFID 114 125 BY SIMILARITY.  
 FT DISULFID 148 193 BY SIMILARITY.  
 FT DISULFID 192 201 BY SIMILARITY.  
 FT DISULFID 224 270 BY SIMILARITY.  
 FT DISULFID 269 277 BY SIMILARITY.  
 FT DISULFID 289 303 BY SIMILARITY.  
 FT DISULFID 302 313 BY SIMILARITY.  
 FT DISULFID 340 385 BY SIMILARITY.  
 FT DISULFID 384 393 BY SIMILARITY.  
 FT DISULFID 416 462 BY SIMILARITY.  
 FT DISULFID 461 472 BY SIMILARITY.  
 FT DISULFID 485 501 BY SIMILARITY.  
 FT DISULFID 500 511 BY SIMILARITY.  
 FT DISULFID 538 583 BY SIMILARITY.  
 FT DISULFID 582 591 BY SIMILARITY.  
 SQ SEQUENCE 608 AA; 68659 MW; 07E629CAC5F60E5F CRC64;  
 Query Match 84.4%; Score 2620; DB 1; Length 608;  
 Best Local Similarity 82.0%; Pred. No. 4e-163;  
 Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;  
 QY 1 DAHSEVAHRPKDLGEENFKALVITAPAYLQQCPFFEDHVKLVNEVTEFAKTCVADSSAE 60  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 25 EAHQSEIAHRRNDLGEHFRGLVLFVAFSQYLQQCPFFEDHVKLVNEVTEFAKGCVDSSAA 84  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLVREPV 120  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 85 NCKSLHELLGDKLCTVASLRDKYGENADCCCKEPPERNECFLOHKDDNPFGQLVTPEA 144  
 QY 121 DMCTAFPHDNEETLKKYLYEIAERHPYFVAPPELLFFAKRYKAAFTCCQADAAKCLLP 180  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 145 DMCTAFPHENQRFGLGKLYEIAERHPYFVAPPELLFYAEYKGVFTTCCEAANDKAACLTP 204  
 QY 181 KLDELRDGKASSAKQRLKCSLQKQGERAFKAMAVARLSQRPFAFAEYKSVKLVTDLTK 240  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 255 KVDALREKVLASSAKERLUKCSLQKQGERAFKANSVARLSQKFPKAFABEISKLVTDLAK 264  
 QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCCCKPLLEKSHCI AEVNDENKPA 300  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 265 LKCCCHGDLLECADRADLAKYICENQDSISTKLKCCCKPVLKSHCI SEVERDESLPA 324  
 QY 301 DLPSLAADPVESKGVCKNYAEKXDFVLFQMFLEYEARHPDYSVYLLRLAKTYETILEK 360  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 325 DLPPLAVDFVEDKEVCKNYQKAKVFGCTFLYEYRRHPESVSVLLRLAKYEAETLEK 384  
 QY 361 CAAADPHECVAKVDFBPKLVEBPONLIVONCELFQOLGEYKFNALLVRYTKVPOVST 420  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 385 CATDDPPACYAHVDFBPKLVEBPONLIVONCELFQOLGEYKFNALLVRYTKVPOVST 444  
 QY 421 PTLVEVSRNLGKVGSKCKHPKAKSCAEDYLSVYVNLGLCVHEKTPVSDRVTKCCTES 480  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 445 PTLVEVSRNLGKVGSKCKHPKAKSCAEDYLSVYVNLGLCVHEKTPVSDRVTKCCTES 504  
 QY 481 LVNRRPCFSALVDEYVVPKFNATFTTFAADCTILSEKERQIKKOTATLVEVVKHKPKAT 540  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 505 LVNRRPCFSALVDEYVVPKFNATFTTFAADCTILSEKERQIKKOTATLVEVVKHKPKAT 564  
 QY 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKKLVAAASQAL 583  
 DB [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||] [|||||]  
 565 BEQLKTVMGDFSGFVDKCAAEDEKAEACFAEBGPKLVAAASQAL 607

## RESULT 4

ALBU\_CANFA STANDARD; PRT; 608 AA.  
 ID ALBU CANFA  
 AC P49822; 077705; Q9TSZ4;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)



QY 541 KEOLKAVMDFAAFVEKCKADKCTCFABEGKKLVAAASQAAL 583  
 D 565 DEQLKTVMDFGAFVEKCKCAENKEGCFSEBGPKLVAQAQAAL 607

## RESULT 5

ALBU HORSE STANDARD; PRT; 607 AA.  
 AC P35737;  
 DT 01-JUN-1994 (Rel. 23, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALBU.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
 RP TISSUE=Liver;  
 RC MEDLINE=93345495; PubMed=8344282;  
 RA Ho J.A., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;  
 RT "X-ray and primary structure of horse serum albumin (Equus caballus)  
 at 0.27-nm resolution."  
 RL Eur. J. Biochem. 215:205-212(1993).  
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 hormones, bilirubin and drugs. Its main function is the regulation  
 of the colloidal osmotic pressure of blood.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Plasma.  
 CC -1- SIMILARITY: BELONGS TO THE ALB/AFP/VDS FAMILY.  
 CC -1- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL: X74045; CAA52194.1; -  
 DR PIR: S34053; ABHOS.  
 DR HSP: P02768; 157B.  
 DR InterPro: IPR000264; Serum albumin.  
 DR Pfam: PF00273; transport\_prot; 3.  
 DR PRINTS: PR00902; SERUMALBUMIN.  
 DR ProDom: PD002496; Serum albumin; 1.  
 DR SMART: SM00103; ALBUMIN; 3.  
 DR PROSITE: PS00212; ALBUMIN; 3.  
 KW Meta-binding; Lipid-binding; Repeat; Signal; Copper.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 607  
 FT DOMAIN 25 204  
 FT DOMAIN 211 396  
 FT DOMAIN 403 594  
 FT META 27 27  
 FT META 27 27  
 FT DISULFID 77 86  
 FT DISULFID 99 115  
 FT DISULFID 114 125  
 FT DISULFID 147 192  
 FT DISULFID 191 200  
 FT DISULFID 223 269  
 FT DISULFID 268 276  
 FT DISULFID 288 302  
 FT DISULFID 301 312  
 FT DISULFID 339 384  
 FT DISULFID 383 392  
 FT DISULFID 415 461

FT DISULFID 460 471  
 FT DISULFID 484 500  
 FT DISULFID 499 510  
 FT DISULFID 537 582  
 FT DISULFID 581 590  
 SQ SEQUENCE 607 AA; 68598 MW; 256F6E810A1B90C5 CRC64;  
 Query Match 79.8%; Score 2475.5; DB 1; Length 607;  
 Best Local Similarity 76.3%; Pred. No. 9.8e-154;  
 Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1;  
 QY 1 DAHSEVAHRPKOLGEENFKALVIAFAQYLOQCPEDHYKLVNEVTEFAKTCVADESAS 60  
 D 25 DTHXSEIAHRFNDLGEKFKGLVLAFAFQYLOQCPEDHYKLVNEVTEFAKTCVADESAS 84  
 QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHODNENLDELVRPEV 120  
 D 85 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHODNENLDELVRPEV 143  
 QY 121 DVMCTAFHDEETFLKKLYEIAARRHPHYFYAPELLFFPKRYKAFTCCQAADKAACLLP 180  
 D 144 DAQCAAFQEDPDKFLGKLYEVAERHPHYFYGPPELLFHAEYKADFTCCPADDKLAELIP 203  
 QY 181 KLDELREGKASSAKQRLKCAISQCKPGERAFKAWAVARLSQRPKAEFAEYKLVITDLTK 240  
 D 204 KLDALKERILLSSAKERLKCSSFQNGERAVKAWAVARLSQRPKAEFAEYKLVITDLTK 263  
 QY 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLECCCKPILLESKSHCAEVENDEMPA 300  
 D 264 VHKCCCHGDLLECADRADLAKYICEHODSISGKLKACCDKPLQKSHCIAEVEDDLPS 323  
 QY 301 DLPSLAADFVESKDVCKRYAEAKDVFLGMFLYEYARRHPDYVSVVLLRLAKTYETILEK 360  
 D 324 DLPALAADFADAEKEICKHYKDAKDVFLGTFLYEYARRHPDYVSVVLLRLAKTYETILEK 383  
 QY 361 CAADPHCEYAKVDEPKLVEEPQNLKQNCLEFQOLGEYKFNALVRYTKVPOVST 420  
 D 394 CAEADPPACVTVFDQPTLVVEPKSLVKKNCDFEEVGYDFONALIVRYTKVPOVST 443  
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCADYLSWLNQLCVLREKTPVSDRVTKCCTES 480  
 D 444 PTLVEIGTLGVKVSRCUKLPESERLPCSENHLALNRLCVLHEKTPVSEKITKCTDS 503  
 QY 481 LVNRRPFSALVEVDETVPKFEFNAETTFHADICTLSEKERQIKKQATLVVELVKKPKAT 540  
 D 504 LAERRPFSALDELDEGVVPKEFKAEFTTFHADICTLPEDEKQIKKQALAEVLVKKPKAT 563  
 QY 541 KEOLKAVMDFAAFVEKCKADKCTCFABEGKKLVAAASQAAL 583  
 D 565 DEQLKTVMDFGAFVEKCKCAENKEGCFSEBGPKLVAQAQAAL 606

## RESULT 6

ALBU BOVIN STANDARD; PRT; 607 AA.  
 ID ALBU BOVIN  
 AC P02769; O82787;  
 DT 21-JUL-1986 (Rel. 31, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor (Allergen Bos d 6).  
 GN ALB.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]\_TaxID=9913;  
 RP SEQUENCE FROM N.A.  
 RA Holowachuk E.W., Stoltenborg J.K., Reed R.G., Peters T. Jr.;  
 RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT THR-214.  
 RC TISSUE=Liver;

RA Barry T., Power S., Gannon F.;  
 R1 Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
 R2 [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A., AND VARIANT THR-214.  
 RA Wu H.T., Huang M.C.;  
 RL "The complete cDNA sequence of bovine serum albumin."  
 RT Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 1-32.  
 RX MEDLINE=80024278; PubMed=488109;  
 RA McGillivray R.T.A., Chung D.W., Davies E.W.;  
 RL "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-  
 RT terminal sequence of prealbumin."  
 RL Eur. J. Biochem. 98:477-485(1979).  
 RN [6]  
 RP SEQUENCE OF 25-424 AND 423-607, AND VARIANT THR-214.  
 RA Brown J.R.;  
 RL "Structure of bovine serum albumin."  
 RT Fed. Proc. 34:591-591(1975).  
 RN [7]  
 RP REVISIONS TO 190-195.  
 RA Brown J.R.;  
 RL Submitted (APR-1975) to the PIR data bank.  
 RN [8]  
 RP SEQUENCE OF 402-433.  
 RX MEDLINE=82023364; PubMed=7283978;  
 RA Reed R.G., Putnam F.W., Peters T. Jr.;  
 RL "Sequence of residues 403-403 of bovine serum albumin."  
 RL Biochem. J. 191:867-868(1980).  
 RN [9]  
 RP SEQUENCE OF 19-28.  
 RX MEDLINE=7134075; PubMed=843354;  
 RA Patterson J.E., Geiler D.M.;  
 RL "Bovine microsomal albumin: amino terminal sequence of bovine  
 RT prealbumin."  
 RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).  
 RN [10]  
 RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.  
 RX MEDLINE=91083649; PubMed=2260975;  
 RA Hirayama K., Akashi S., Furuya M., Fukuhara K.-I.;  
 RL "Rapid confirmation and revision of the primary structure of bovine  
 RT serum albumin by EIMS and Frit-FAB LC/MS."  
 RL Biochem. Biophys. Res. Commun. 173:639-646(1990).  
 RN [11]  
 RP SEQUENCE OF 25-41.  
 RX MEDLINE=88267456; PubMed=3389500;  
 RA Hsieh J.C., Lin F.P., Tam M.F.;  
 RL "Electroblotting onto glass-fiber filter from an analytical  
 RT isoelectrofocusing gel: a preparative method for isolating proteins  
 RL for N-terminal microsequencing."  
 RL Anal. Biochem. 170:11-8(1988).  
 RN [12]  
 RP SEQUENCE OF 437-451.  
 RA Vilbois F.;  
 RL Submitted (AUG-1998) to the SWISS-PROT data bank.  
 RN [13]  
 RP DISULFIDE BONDS.  
 RA Brown J.R.;  
 RL "Structure of serum albumin: disulfide bridges."  
 RT Fed. Proc. 33:1389-1389(1974).  
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloid osmotic pressure of blood.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Plasma.  
 CC -1- SIMILARITY: BELONGS TO THE ALB/AF/VDL FAMILY.  
 CC -1- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X73993; AA051411.1; -;  
 DR EMBL; X58989; CAA41735.1; -;  
 DR EMBL; Y17769; CAA76847.1; -;  
 DR EMBL; AF542068; AAM17824.1; -;  
 DR HSSP; P02768; 1E7B;  
 DR InterPro; IPR000264; Serum\_albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR ProTis; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;  
 KW Peptidomorphism.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 607 SERUM ALBUMIN.  
 FT DOMAIN 25 204 ALBUMIN 1.  
 FT DOMAIN 211 396 ALBUMIN 2.  
 FT DOMAIN 403 594 ALBUMIN 3.  
 FT METAL 27 27 COPPER (BY SIMILARITY).  
 FT DISULFID 77 86  
 FT DISULFID 99 115  
 FT DISULFID 114 125  
 FT DISULFID 147 192  
 FT DISULFID 191 200  
 FT DISULFID 223 269  
 FT DISULFID 268 276  
 FT DISULFID 288 302  
 FT DISULFID 301 312  
 FT DISULFID 339 384  
 FT DISULFID 383 392  
 FT DISULFID 415 461  
 FT DISULFID 460 471  
 FT DISULFID 484 500  
 FT DISULFID 499 510  
 FT DISULFID 537 582  
 FT DISULFID 581 590  
 FT VARIANT 214 214  
 FT CONFLICT 302 302 A -> T.  
 FT CONFLICT 304 305 C -> K (IN REF. 6).  
 FT CONFLICT 324 324 KP -> PC (IN REF. 6).  
 FT CONFLICT 324 324 N -> D (IN REF. 6).  
 FT CONFLICT 394 395 ST -> TS (IN REF. 6).  
 FT CONFLICT 437 437 K -> R (IN REF. 12).  
 FT CONFLICT 493 494 SE -> ES (IN REF. 6).  
 SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;  
 Query Match 79.0%; Score 2450.5; DB 1; Length 607;  
 Best Local Similarity 75.8%; Pred. No. 4,1e-152;  
 Matches 442; Conservative 71; Mismatches 69; Indels 1; Gaps 2;  
 QY 1 DAHKSVAHRFKDLGEEHFKALVLIAPAYLQCPEDHVKLVNEVTEPAKTCVADESAE 60  
 DB 25 DTHKSEIAHFRFKDLGEEHFKGLVLIAPFQYLOQCPEDHVKLVNELTEPAKTCVADESHA 84  
 QY 61 NCDKSLATLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120  
 DB 85 GCEKSLJTLFGDELCKVASRETYGMDCCCKQEPERNECFLSHKDDSPDLPLKLP 143  
 QY 121 DVMCTAFTHDNEETFLKKLYEIAARRHPFYFVAPELJFFAKYKAAFTCCQAAADKAACLLP 180  
 DB 144 NTLCDPEKADKKFKWGLYELIARRHPFYFVAPELLYYANKYNGVFQCCQADKGAQLLP 203  
 QY 161 KLDELDEGVASSAKQLKLCASLOKGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240

Db 204 KIETVREKVLASSARCKLRCSAQKFGERALKAWSVARLSQKFPKAEFVEVTKLVDTLTK 263

QY 241 VHTTECHGDLLECCADBRADLAKVYCENQDSISSKLEKCEKPKLLEKSHCIAEVENDEMPA 300

Db 264 VHKCECHGDLLECCADBRADLAKVYCENQDSISSKLEKCEKPKLLEKSHCIAEVEKDAPE 323

QY 301 DPLSLAADFVESKVCNKYAAEKDFLGMFLYAYARRHPDYSVVLRLRIAKTYETTLK 360

Db 324 NLPLTADPAEDKVCNKYQENKDAFLGSLFYYSRRHPYAVSVJRLAKYEATLEEC 383

QY 361 CAADAPHECYAKVDFDEKPVVEBPQNDLQKNCBELFQJAEYKFNQALJYVYTKKVPQVST 420

Db 384 CAKDQPHACYSTVDFKLKLVDEPNLQKNCDFEKLGEYFQNALJYVYTRKVPQVST 443

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLVSUVLNGLCVLHHEKTPYSDRVTKCCTES 480

Db 444 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLVSUVLNGLCVLHHEKTPYSDRVTKCCTES 503

QY 481 LVNRRPFCFSALEVDITYVPKFEFNAETTFHADICTLSEKERQTKKOTALVELVGHKPKAT 540

Db 504 LVNRRPFCFSALEVDITYVPKFEFNAETTFHADICTLSEKERQTKKOTALVELVGHKPKAT 563

QY 541 KEQLKAVMDDFAFVPEKCCKADDKETCFABEGKLVVAASQAL 593

Db 564 EQQLTKVMENFAVFDKCAADDKAEACFAVEGPKLVVSTQAL 606

## RESULT 7

ALBU\_RABIT STANDARD; PRT; 608 AA.

AC P49065;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Serum albumin precursor.

GN ALB.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eucleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI\_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=New Zealand white; TISSUE=Liver;

RA Sheffield W.P., Syed S., Schuyler P.C.;

RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good

CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,

CC hormones, bilirubin and drugs. Its main function is the regulation

CC of the colloidal osmotic pressure of blood.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Plasma.

CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.

CC -!- SIMILARITY: Contains 3 albumin domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

CC EMBL; U18344; AAB58347.1; -

DR HSSP; P02768; IE75.

DR InterPro: IPR000264; Serum albumin.

DR Pfam: PF00273; transport prot; 3.

DR PRINTS: PR00802; SERJALBUMIN.

DR ProDom: PD002486; Serum albumin; 1.

DR SMART: SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.

FT SIGNAL 1 18 BY SIMILARITY.

FT PROPEP 19 24 BY SIMILARITY.

FT CHAIN 25 608 SERUM ALBUMIN.

FT DOKAIN 25 205 ALBUMIN 1.

FT DOKAIN 212 397 ALBUMIN 2.

FT DOKAIN 404 595 ALBUMIN 3.

FT METAL 27 27 COPPER.

FT DISULFID 77 86 BY SIMILARITY.

FT DISULFID 99 115 BY SIMILARITY.

FT DISULFID 114 125 BY SIMILARITY.

FT DISULFID 148 193 BY SIMILARITY.

FT DISULFID 192 201 BY SIMILARITY.

FT DISULFID 224 270 BY SIMILARITY.

FT DISULFID 269 277 BY SIMILARITY.

FT DISULFID 289 303 BY SIMILARITY.

FT DISULFID 302 313 BY SIMILARITY.

FT DISULFID 340 385 BY SIMILARITY.

FT DISULFID 384 393 BY SIMILARITY.

FT DISULFID 416 462 BY SIMILARITY.

FT DISULFID 461 472 BY SIMILARITY.

FT DISULFID 485 501 BY SIMILARITY.

FT DISULFID 500 511 BY SIMILARITY.

FT DISULFID 538 583 BY SIMILARITY.

FT DISULFID 582 591 BY SIMILARITY.

SQ SEQUENCE 608 AA; CF5E92647AAFE9A2 CRC64;

Query Match 78.8%; Score 2446; DB 1; Length 608;

Best Local Similarity 74.1%; Pred. No. 8,1e-152;

Matches 433; Conservative 77; Mismatches 74; Indels 0; Gaps 3;

QY 1 DAHKEVAHRRFKDLGSENFKAALVJAFAYLOQCFEDHVKLVNEVTEFAKTCVADSEAE 60

DB 25 EAHKSEIARFNDVGEHFHIGJLVITFSYLOKCPYEEHAKLVKEVTDLAKACVADSEAA 84

QY 61 NCDKSLHCTLFQDKLCTVATLRETYGEMADCCAKPEPERNECFLOHKDDNPNLPLRLVPEV 120

DB 85 NCDKSLHDIQDKICALPSLRDTYGVADCCKEKEPERNECFLLHKKDKDPOLPFAPEA 144

QY 121 DVMCTAFHNDNEETFLKLYLYEIAIRRHYPFYAPPELLFFAKRYKAAFTCCQADKACLLP 180

DB 145 DVLCKAFHDEKAFEGHYLYEVARRHPYFYAPPELLLYAOKYKAILTECCAAAKGACLT 204

QY 181 KLDELREGKASSAKQRLKCLASLOKFGERAFKAVARLSORFPKAEFAEVSCKVTDLT 240

DB 205 KLDALGKSLISAQGERLURCASIOKFGDRAYKAWALVRLSORFPKADFTDISKIVTDLT 264

QY 241 VRTCECHGDLLECCADBRADLAKVYCENQDSISSKLEKCEKPKLLEKSHCIAEVENDEMPA 300

DB 265 VHKCECHGDLLECCADBRADLAKVYCENQDSISSKLEKCEKPKLLEKSHCIAEVEKDAPE 324

QY 301 DPLSLAADFVESKVCNKYAAEKDFLGMFLYAYARRHPDYSVVLRLRIAKTYETTLK 360

DB 325 GJPAVAEEFVEDKQVCKNYEBAKDLFLGKPYEYSRHPDYSVVLRLRLGKAYEATLK 384

QY 361 CAADAPHECYAKVDFDEKPVVEBPQNDLQKNCBELFQJAEYKFNQALJYVYTRKVPQVST 420

DB 385 CATDDPHACYAKVDFDEKPVVEBPQNDLQKNCBELFQJAEYKFNQALJYVYTRKVPQVST 444

QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLVSUVLNGLCVLHHEKTPYSDRVTKCCTES 480

DB 445 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLVSUVLNGLCVLHHEKTPYSDRVTKCCTES 504

QY 481 LVNRRPFCFSALEVDITYVPKFEFNAETTFHADICTLSEKERQTKKOTALVELVGHKPKAT 540

DB 505 LVNRRPFCFSALEVDITYVPKFEFNAETTFHADICTLSEKERQTKKOTALVELVGHKPKAT 564

QY 541 KEQLKAVMDDFAFVPEKCCKADDKETCFABEGKLVVAASQALG 584

DB 565 NDQLTKVMGEFTALDDKCCSAEDKEACFAVEGPKLVVSTQALG 608

## RESULT 8

ALBU SHEEP

ID ALBU SHEEP PRT; 607 AA.

AC P14639;

DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serum albumin precursor.  
GN ALB.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=9009888; PubMed=2602150;  
RA Brown W.M., Dziegielewska K.M., Foreman R.C., Saunders N.R.;  
RT "Nucleotide and deduced amino acid sequence of sheep serum albumin.";  
RL Nucleic Acids Res. 17:10495-10495(1989).  
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
CC hormones, bilirubin and drugs. Its main function is the regulation  
CC of the colloidal osmotic pressure of blood.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Plasma.  
CC -!- SIMILARITY: BELONGS TO THE ALB/AFI/VBS FAMILY.  
CC -!- SIMILARITY: Contains 3 albumin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).  
CC  
DR EYBL; X17055; CAA34903.1; -  
DR PIR; S06936; ABSHS.  
DR HSSP; P02768; 1E7E.  
DR InterPro; IPR000264; Serum\_albumin.  
DR Pfam; PF00273; transport\_prot; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PD002486; Serum\_albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT SIGNAL 1 18 BY SIMILARITY.  
FT PROPEL 19 24 BY SIMILARITY.  
FT CHAIN 25 607 SERUM ALBUMIN.  
FT DOMAIN 25 204 ALBUMIN 1.  
FT DOMAIN 211 396 ALBUMIN 2.  
FT DOMAIN 403 594 ALBUMIN 3.  
FT METAL 27 27 COPPER (BY SIMILARITY).  
FT DISULFID 77 86 BY SIMILARITY.  
FT DISULFID 99 115 BY SIMILARITY.  
FT DISULFID 114 125 BY SIMILARITY.  
FT DISULFID 147 192 BY SIMILARITY.  
FT DISULFID 191 200 BY SIMILARITY.  
FT DISULFID 223 269 BY SIMILARITY.  
FT DISULFID 268 276 BY SIMILARITY.  
FT DISULFID 288 302 BY SIMILARITY.  
FT DISULFID 301 312 BY SIMILARITY.  
FT DISULFID 339 384 BY SIMILARITY.  
FT DISULFID 383 392 BY SIMILARITY.  
FT DISULFID 415 461 BY SIMILARITY.  
FT DISULFID 460 471 BY SIMILARITY.  
FT DISULFID 484 500 BY SIMILARITY.  
FT DISULFID 499 510 BY SIMILARITY.  
FT DISULFID 537 582 BY SIMILARITY.  
FT DISULFID 581 590 BY SIMILARITY.  
SQ SEQUENCE 607 AA; 69198 MW; 84979AB7FB86596 CRC64;  
Query Match 78.4%; Score 2432.5; DB 1; Length 607;  
Best Local Similarity 75.0%; Pred. No. 6.1e-151;  
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

QY 1 DAHKGEVAHRFKDLGEENFKALVLJAFAYLQQCFEDPHVKLVNEVTEFAKTQVADSAAE 60  
DB 25 DTHKSEIAHRFNDLGEENFQGLVLJAFASQYLQQCFEDPHVKLVKELTEFAKTQVADSAAE 84  
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDNPRLPLRVREV 120  
DB 85 GCDKSLHTLFGDELCKVATLRETYGDMADCCKEGPERNECFLNKKDQSPDLPKL-KPEP 143  
QY 121 DVMCTAFHDNEETFLKKYLYEIAHRHPYFYAPPELLFFAKRYKAAAFTECCQAADKAACLLP 180  
DB 144 DTLCAEFKADKKFWGKLYEVARRHPYFYAPPELLYYANKYNGVFOECCQAEDKGCALLP 203  
QY 181 KLDELRLDECKASSAKORLKCAILOKFGESAPKANAVARLSQRFPAKAEFAEYSKLVTLTK 240  
DB 204 KIDAMREKVLASSARORLCASIQKFGERALKAMSVARLSQKFPKACFTDVTK-VTDLTK 263  
QY 241 VTECHGGLLECADRADLAKYICENODSISSKLKECCXPLLEKSHSCIAEVENDEMPA 300  
DB 264 VHKCOCHGLLSCADDRADLAKYICDHCDALUSSKLKECCDKPVLKSHSCIAEVDKDAVPE 323  
QY 301 DLPLSLAADFVSKDYCKNYBAKQVFLQMFYEVARRHPDYVSVLLRLAKTYETTLTK 360  
DB 324 NLPLPLTADPAEDKCYKQYQEAQVFLGSLFYEXSRHPYAVSVLLSLAKEYEATLEDC 383  
QY 361 CAAADPHECYAKVDFEKPVEEPQNLIKONCELFEOLGEYKFGONALVRYTKVPOVST 420  
DB 384 CAKEDPHACYATVFDKLKHLVDPEQNLIKONCELFEKGEYFGONALVRYTKRQVST 443  
QY 421 PTLVEVSRNLGKVGSKCKHPKAPKAPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTTES 480  
DB 444 PTLVEISRLGKVGTKKCAKPESEMPCTEDYLSILNKLCLVLEKTPVSEKVTKCTTES 503  
QY 481 LVNRRPCFSALEVDYTPYKFNATFTFHADICTLSEKERCIKKQTALVELVGHKPKAT 540  
DB 504 LVNRRPCFSOLTLDETYYKPFDEKFFTFHADICTLFDTEKQIKKQTALVELVGHKPKAT 563  
QY 541 KEQLKAVMDQDFAAFVEKCKKADDKETCTFAEEGKKLVAAASQAAL 583  
DB 564 DEOLKTVMENFVAFVDKCCAADKSGCFVLEGPKVAVSTQAAAL 606  
RESULT 9  
ALBU RAT  
ID ALBU RAT STANDARD; PRT; 608 AA.  
AC P02770; P11382;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP);  
GN ALB.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=81223722; PubMed=7017712;  
RA Sargent T.D., Yang W., Bonner J.;  
RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).  
RN [2]  
RP SEQUENCE OF 1-38, AND PROCESSING.  
RX MEDLINE=77249657; PubMed=893447;  
RA Strauss A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;  
RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-  
RT piece. Analysis of the direct translation product of albumin  
RT messenger RNA.";  
RL J. Biol. Chem. 252:6846-6855(1977).  
RN [3]  
RP SEQUENCE OF 25-222.  
RX MEDLINE=78109429; PubMed=564345;  
RA Isemura S., Ikenaka T.;

RT "Amino acid sequences of fragments I and II obtained by cyanogen  
 RT bromide cleavage of rat serum albumin.";  
 RL J. Biochem. 83:35-48(1978).  
 RN [4]  
 RP SEQUENCE OF 223-288 AND 572-608.  
 RX MEDLINE=76260153; PubMed=956149;  
 RA Ikenaka T.;  
 RA "Fragmentation of rat serum albumin by cyanogen bromide cleavage and  
 RT the amino acid sequences of four fragments.";  
 RL J. Biochem. 79:1183-1196(1976).  
 RN [5]  
 RP SEQUENCE CF 168-174.  
 RX TISSUE=Plasma;  
 RA MEDLINE=87194805; PubMed=2437117;  
 RA Carraway R.E., Mitra S.P., Cochrane D.E.;  
 RT "Structure of a biologically active neurotensin-related peptide  
 RT obtained from pepsin-treated albumin(s).";  
 RL J. Biol. Chem. 262:5968-5973(1987).  
 RN [6]  
 RP COPPER-BINDING.  
 RX MEDLINE=79001617; PubMed=80265;  
 RA Aoyagi Y., Ikenaka T., Ichida F.;  
 RT "Copper(II)-binding ability of human alpha-fetoprotein.";  
 RL Cancer Res. 38:3483-3486(1978).  
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloid osmotic pressure of blood.  
 CC -1- FUNCTION: NRP REGULATES FAT DIGESTION, LIPID ABSORPTION, AND  
 CC BLOOD FLOW (POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Plasma.  
 CC -1- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.  
 CC -1- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; V01222; CAA24532.1; -;  
 DR PIR; A93872; ABRTS.  
 DR HSSP; P02768; IEBH.  
 DR InterPro; IPSC00264; Serum albumin.  
 DR Pfam; PF00273; transport\_pro\_3.  
 DR PRINTS; PR08802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum\_albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 FT SIGNAL 1 18  
 FT PROPEP 19 24  
 FT CHAIN 25 608 SERUM ALBUMIN.  
 FT PEPTIDE 166 174 NEUTROTENSIN-RELATED PEPTIDE.  
 FT DOMAIN 25 205 ALBUMIN 1.  
 FT DOMAIN 212 397 ALBUMIN 2.  
 FT DOMAIN 404 595 ALBUMIN 3.  
 FT METAL 27 27 COPPER.  
 FT DISULFID 77 86 BY SIMILARITY.  
 FT DISULFID 99 115 BY SIMILARITY.  
 FT DISULFID 114 125 BY SIMILARITY.  
 FT DISULFID 148 193 BY SIMILARITY.  
 FT DISULFID 192 201 BY SIMILARITY.  
 FT DISULFID 224 270 BY SIMILARITY.  
 FT DISULFID 269 277 BY SIMILARITY.  
 FT DISULFID 289 303 BY SIMILARITY.  
 FT DISULFID 302 313 BY SIMILARITY.  
 FT DISULFID 340 385 BY SIMILARITY.  
 FT DISULFID 384 393 BY SIMILARITY.  
 FT DISULFID 416 462 BY SIMILARITY.

FT	DISULFID	461	472		BY SIMILARITY.
FT	DISULFID	485	501		BY SIMILARITY.
FT	DISULFID	500	511		BY SIMILARITY.
FT	DISULFID	538	583		BY SIMILARITY.
FT	DISULFID	582	591		BY SIMILARITY.
FT	VARIANT	262			V -> L.
FT	CONFLICT	174			Y -> S (IN REF. 5).
SQ	SEQUENCE	608 AA;	687-8 MW;	538497A28241AB7	CRC64;

Query Match 78.2%; Score 2426; DB 1; Length 608;  
 Best Local Similarity 73.4%; Pred. No. 1.6e-150;  
 Matches 428; Conservative 82; Mismatches 73; Indels 0;

QY	1	DAKSEVAHREFKDLGEENFKALVLIATFAQYQQCFPFHDHVKLVNEVTEFAKTCVADSAE	60
DB	25	BAHKSEIAHREFKDLGEGHFKGLVLIATFSQYLOKFPYEEHKLKVGVIDTFAKTCVADSAE	84
QY	61	NCDKSLHTLFGDKLCTVATLRETYGEVADCCAKGEPERNECFLOHKDNDNPLPLVPEV	120
DB	85	NCDKSIHTLFGDKLCAIPKLRDNYGELADCCAKGEPERNECFLOHKDNDNPLPLVPEV	144
QY	121	DMCTAFHDNEETFKKYLYEYARRHPYFAPPELLFAKRYKAATFECQADGAACLLP	180
DB	145	EMCTSPQENTSTFGHYLHEVARHPYFAPPELLYAEKYNELVTQCTESDKAACLLP	204
QY	191	KIDELRDEGKASSAKQRLKCSAQKGERAFKAWAVARLSQRFPAKAEVSKLVTDLT	240
DB	205	KLDVKEKALVAAVRQRMKCSMQRFGERAFKAWAVARMSQRFPAEFAETKLTADVT	264
QY	241	VHTECHGDLLECADDRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMP	300
DB	265	INKECHGDLLECADDRADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVENDEMP	324
QY	301	DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYVLLRLAKTYETILEK	360
DB	325	DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYVLLRLAKTYETILEK	384
QY	361	CAAADPHCYAKVDFEKPVEEPQNLKONCEFEQLGEYKFNALLVRYTKVPQVST	420
DB	385	CAEGDPACYGVLAEPQVLEEPAKLVKTNKELVEKLGEGFQNAVLRVYTKQAPQVST	444
QY	421	PTLVEVSNLKVGSKCKCKHPEAKRMPQAEYLSVLNQLVLEKTPVSDRVTKCCTES	480
DB	445	PTLVEVSNLKVGSKCKCKHPEAKRMPQAEYLSVLNQLVLEKTPVSDRVTKCCTES	504
QY	481	LVNRPFCFSALEVDSTVPKPFNAETFTFHADICTLSEKESQIKKOTALVELVKHKKAT	540
DB	505	LVNRPFCFSALEVDSTVPKPFNAETFTFHADICTLSEKESQIKKOTALVELVKHKKAT	564

RESULT 10  
 ID ALBU PIG STANDARD; PR7; 605 AA.  
 AC P08635; Q29018;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE .Serum albumin precursor (Fragment).  
 GN ALB.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver.  
 RX MEDLINE=89016582; PubMed=3174440;  
 RA Baldwin G.S., Weinstock J.;  
 RT "Nucleotide sequence of porcine liver albumin.";

RL Nucleic Acids Res. 16:9045-9045(1989).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 DR EMBL; X12422; CAA30970.1; -;  
 DR EMBL; M36787; CAA30988.1; -;  
 DR PIR; S01382; ABOS.  
 DR HSSP; P02768; 1E7H.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR PRODOM; PD002486; Serum albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 16 BY SIMILARITY.  
 FT PROPEP 17 22 BY SIMILARITY.  
 FT CHAIN 23 605 SERUM ALBUMIN.  
 FT DOKAIN 23 202 ALBUMIN 1.  
 FT DOKAIN 209 394 ALBUMIN 2.  
 FT DOKAIN 401 592 ALBUMIN 3.  
 FT METAL 31 31 COPPER (BY SIMILARITY).  
 FT DISULFD 31 84 BY SIMILARITY.  
 FT DISULFD 97 123 BY SIMILARITY.  
 FT DISULFD 112 123 BY SIMILARITY.  
 FT DISULFD 145 190 BY SIMILARITY.  
 FT DISULFD 199 198 BY SIMILARITY.  
 FT DISULFD 221 267 BY SIMILARITY.  
 FT DISULFD 266 274 BY SIMILARITY.  
 FT DISULFD 286 300 BY SIMILARITY.  
 FT DISULFD 299 310 BY SIMILARITY.  
 FT DISULFD 337 362 BY SIMILARITY.  
 FT DISULFD 381 390 BY SIMILARITY.  
 FT DISULFD 413 459 BY SIMILARITY.  
 FT DISULFD 458 469 BY SIMILARITY.  
 FT DISULFD 482 498 BY SIMILARITY.  
 FT DISULFD 497 508 BY SIMILARITY.  
 FT DISULFD 535 580 BY SIMILARITY.  
 FT DISULFD 579 588 BY SIMILARITY.  
 FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).  
 SQ SEQUENCE 605 AA; 69410 MW; 35556B0DD1A1F4FF CRC64;

Query Match  
 Best Local Similarity 77.7%; Score 2411.5; DB 1; Length 605;  
 Matches 438; Conservative 67; Mismatches 70; Indels 1; Gaps 1;  
 QY 1 DAHKEVAHRFKDGEENFKALVIAFAQYLCQCFEDHVKLVNVEVTEFAKTCVADESAE 60  
 DB 23 DTJKSEIAHRPNKDLGEQYFKGLVIAFASOHLQCCQPYBEHVKLVREVEVTEFAKTCVADESAE 92  
 QY 61 NCDKSLHTLFGDKCLVTATLSTGEMACCAKOEBSNECFLOHKQDNPNLRLVSEPV 120  
 DB 83 NCDKSLHTLFGDKCA:PSLREHYOCLADCKEKEPERNECFLOHKQDNPNLRLVSEPV 141  
 QY 121 DVMTAFHNEETFLKKYLYEIAHRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180  
 DB 142 VALCADFQDEQKFGKYGKLYEIAHRHPYFYAPPELLYVYAIYKCVSECCQAADKAAACLLP 201

QY 181 KLDSEIRDEGKASSAKQRLKCSAQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTIX 240  
 DB 202 KIEHLREKVLTSAAKQRLKCSAQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTIX 261  
 QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISKLKCECKPLEKSHCIAEVENDEMPA 300  
 DB 262 VKCCCHGDLLECADRADLAKYICENQDTISTKLKCCDKPLEKSHCIAEAKRDELEPA 321  
 QY 301 DLPSLAADPFVESKDVCKYABAKDVFLGMFLYFYARRHPDYSVLLRLAKTYETTLEK 360  
 DB 322 DLNPLEHDFVEDKEVCNKYAEKDVFLGFLTYEYRRHPDYSVLLRLIAK:YEATLEDC 381  
 QY 361 CAADHPCEYAKYDFEKP:VESEPNLIKONCE:FECLGEYKTONALLVRYTKKVPQVST 420  
 DB 382 CAKEPPACVATVDFKQPPVDSEPNLIKONCELFKJGEGYGFQNALIVRYTKKVPQVST 441  
 QY 421 PTLVEVSRLNGKSGKCKHPEAKRYPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480  
 DB 442 PTLVEVARKLGLVGRCCCKRPEERLSCAEDYLSLV:NRILCVLHEKTPVSEKVTCTES 501  
 QY 481 LVNRRPCFSA:LEVDETYVPKFNARPTTFHADICTLSEKERQIKKQTALVELVHKPKAT 540  
 DB 502 LVNRRPCFSALTPTDITYKPEFVEGTTFHADJCTLPEDSKQIKKQTALVELVHKPKAT 561  
 QY 541 KEOLKAVNDPFAAFVEKCKKADDKETCFABEGKLV 576  
 DB 562 ZEQLRTV:GNFAAFVQKCAAPDHEACFAVEGPKFV 597  
 RESULT 11  
 ALBU\_MERUN  
 ID ALBU\_MERUN STANDARD; PRT: 609 AA.  
 AC O35090;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB.  
 OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil);  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;  
 OC Meriones.  
 CX NCBI\_TaxID=10047;  
 RN 1;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MGS IDR; TISSUE=Liver;  
 RX MEDLINE=9811663; PubMed=9455485;  
 RA Yoshida K., Seto-Onshima A., Sinozawa H.;  
 RT "Sequencing of cDNA encoding serum albumin and its extrahepatic  
 RT synthesis in the Mongolian gerbil, Meriones unguiculatus.";  
 RL DNA Res. 4:351-354(1997).  
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Plasma.  
 CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 DR EMBL; AB006197; BAA21765.1; -;  
 DR PIR; JCS838; JCS838.  
 DR HSSP; P02768; 1E7H.  
 DR InterPro; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transport\_prot; 3.



CC -- TISSUE SPECIFICITY: Plasma.  
 CC -- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.  
 CC -- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AJ011413; CAA09617.1; --  
 CC DR EMBL: M1611; AAA37190.1; --  
 CC DR EMBL: X13063; CAA31458.1; --  
 CC DR EMBL: AK010025; BAB26653.1; --  
 CC DR HIR: A35139; A05139.  
 CC DR HIR: P02768; LEB7.  
 CC DR PIR: P02768; LEB7.  
 CC DR SWISS-2DPAGE: P07724; MOUSE.  
 CC DR MGD: MGI:87591; Alb1.  
 CC DR InterPro: IP8000264; Serum albumin.  
 CC DR Pfam: PF00273; transprot\_prot; 3.  
 CC DR PRINTS: PR08002; SERUMALBUMIN.  
 CC DR ProDom: PD002486; Serum\_albumin; 1.  
 CC DR SMART: SM00103; ALBUYN; 3.  
 CC DR PROSITE: PS00212; ALBUMIN; 3.  
 CC KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
 CC FT SIGNAL: 1 18  
 CC FT PROPEP: 19 24  
 CC FT CHAIN: 25 608  
 CC FT DOMAIN: 25 205  
 CC FT DOMAIN: 212 397  
 CC FT DOMAIN: 404 595  
 CC FT METAL: 27 27  
 CC FT DISULFID: 77 86  
 CC FT DISULFID: 99 115  
 CC FT DISULFID: 114 125  
 CC FT DISULFID: 148 193  
 CC FT DISULFID: 192 201  
 CC FT DISULFID: 224 270  
 CC FT DISULFID: 269 277  
 CC FT DISULFID: 289 303  
 CC FT DISULFID: 302 313  
 CC FT DISULFID: 340 385  
 CC FT DISULFID: 384 393  
 CC FT DISULFID: 416 462  
 CC FT DISULFID: 461 472  
 CC FT DISULFID: 500 511  
 CC FT DISULFID: 538 583  
 CC FT DISULFID: 582 591  
 CC FT CONFLICT: 27 27 H -> D (IN REF. 5).  
 CC FT CONFLICT: 33 33 H -> D (IN REF. 5).  
 CC FT CONFLICT: 41 41 Q -> I (IN REF. 5).  
 CC SQ SEQUENCE 608 AA; 68692 MW; 292F7C7EED3A61B4 CRC64;  
 Query Match 76.6%; Score 2378; DB 1; Length 608;  
 Best Local Similarity 72.4%; Pred. No. 2.1e-147;  
 Matches 422; Conservative 80; Mismatches 91; Indels 0; Gaps 0;  
 QY 1 DAHSEVAHFKDGLGKGFALVIAFAQVLOQCPEDHVKLVNEVTEFAKTCVADESAA 60  
 Db 25 EAHKSEIAHRYNDLGEHFGKGLVLIASFQVLCSDYDEAKLVQEVTFDAKTCVADESAA 84  
 QY 61 NCDKSLHTLFGDKLCTVATLRYTGEMADCCAKQEPERNECFQHKDDNPRLVPRPV 120  
 Db 85 NCDKSLHTLFGDKLCAIPNLRENYGELADCCCTQEPERNECFQHKDDNPRLVPRPEA 144  
 QY 121 DVNCTAFHNEETFLKKYLVEIRRPYFYVAPELLFFAKRYKGAFTCCQCAKCAACILP 180  
 Db 145 EAMCTSFKNPTTFMGHYLHEVARRHPYFYVAPELLVYAEQNEILTQCAEAKKSCILTP 204  
 QY 181 KLDELDEGKASSAKQRLKCAKSLQKFGERAFKAWAVARLSQRFPAKFAEVSKLVTDLT 240

Db 205 KLGKVKALVSSVRQMKCSSMOKFGERAFKAWAVARLSQTFPNADFAETTKUATDLTK 264  
 QY 241 VHTCCGGDLLECADRADLAKYICENQDSISSKKECCCKPLLEKSHCIAEVENDEMPA 300  
 Db 265 VNKCCGGDLLECADRADLAKYICENQDSISSKKECCCKPLLEKSHCIAEVENDEMPA 324  
 QY 301 DLPSLAADFVSKDVKYNAEAKDVFLGMFLYEVARRHPDYVSVLLRLAKTYETTLKCC 360  
 Db 325 DLPAIAADFVEDQEVCKYNAEAKDVFLGMFLYEVARRHPDYVSVLLRLAKTYETTLKCC 384  
 QY 361 CAADAPHECYAKVDFEKPVEBPONLIKONCELFEGLEGYKFFONA:LVRYTKVPOVST 420  
 Db 385 CAENPPACVGTVLAEFQPLVEEPONLVKNCDDLYEKLGEYGFONALVRYTKVPOVST 444  
 QY 421 PTLVEVSRNLGKSGCKCKHPEAKMPCAEADYLSVWLNQLCVLHETPVSDRVTKCCTES 480  
 Db 445 PTLVEAARNLGRVGTCKCTLPEDQRLPCVEDYLSAILNRVCLLHETPVSERVTKCCSGS 504  
 QY 481 LVNRPFCFSALEVDVETVVPKEFNAETTFHADICTLSEKEROIKKQTALVELYKHKPKAT 540  
 Db 505 LVERRPCFSALTVDVETVVPKEFNAETTFHADICTLSEKEROIKKQTALVELYKHKPKAT 564  
 QY 541 KEQLKAVMDQFAAFVEKCCCKADDDKCTCFABEGKGLVAASOAAAL 583  
 Db 565 AEQLKXTVMDDQFAQLDTCCKAADKCTCFSTEGPNLVTRCKDAL 607  
 RESULT 13  
 ALBU CHICK STANDARD; ERT; 615 AA.  
 ID ALBU CHICK  
 AC P19:21;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Serum albumin precursor.  
 GN ALB.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI TaxID=9031;  
 RN [1] \_TaxID=9031;  
 RC SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Cassidy A., Salkild C.K., Baverstock P., Wallace J.C.;  
 RJ Submitted (JUL-1991) to the EMBL/GenBank/CCDB databases.  
 RN [2]  
 RP SEQUENCE OF 1-28 FROM N.A.  
 RX MEDLINE=83161037; PubMed=6187737;  
 RA Hache R.J.G., Wiskocil R., Vasa M., Roy R.N., Lau P.C.K., Deeley R.G.;  
 RT "The 5' noncoding and flanking regions of the avian very low density  
 RT apolipoprotein II and serum albumin genes. Homologies with the egg  
 RT white protein genes.";  
 RL J. Biol. Chem. 258:4556-4564(1983).  
 RN [3]  
 RP SEQUENCE OF 19-30.  
 RX MEDLINE=78019943; PubMed=911327;  
 RA Rosen A.M., Geller D.M.;  
 RT "Chicken microsomal albumin: amino terminal sequence of chicken  
 RT proalbumin.";  
 RL Biochem. Biophys. Res. Commun. 78:1060-1066(1977).  
 CC -- FUNCTION: Serum albumin, the main protein of plasma, has a good  
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
 CC hormones, bilirubin and drugs. Its main function is the regulation  
 CC of the colloidal osmotic pressure of blood.  
 CC -- SUBCELLULAR LOCATION: Secreted.  
 CC -- TISSUE SPECIFICITY: Plasma.  
 CC -- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.  
 CC -- SIMILARITY: Contains 3 albumin domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

```

CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/)
CC CC or send an email to license@isb-sib.ch.
CC CC -----
DR DR EMBL: X60688; CAA43098.1; -
DR DR EMBL: V00381; CAA23680.1; -
DR DR PIR: S15571; ABCSH.
DR DR HSP: P02768; IE7B.
DR DR InterPro: IPR000264; Serum_albumin.
DR DR Pfam: PF00273; transport_prot; 3.
DR DR PRINTS: P008002; SERUMALBUMIN.
DR DR ProDom: PD002486; Serum_albumin; 1.
DR DR SMART: SM00103; ALBUMIN; 3.
DR DR PROSITE: PS00212; ALBUMIN; 3.
KW Meta-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 23
FT CHAIN 24 615 SERUM ALBUMIN.
FT DOMAIN 24 209 ALBUMIN 1.
FT DOMAIN 216 401 ALBUMIN 2.
FT DOMAIN 403 599 ALBUMIN 3.
FT METAL 30 30 COPPER (BY SIMILARITY).
FT DISULFID 80 89 BY SIMILARITY.
FT DISULFID 102 118 BY SIMILARITY.
FT DISULFID 117 128 BY SIMILARITY.
FT DISULFID 152 197 BY SIMILARITY.
FT DISULFID 196 205 BY SIMILARITY.
FT DISULFID 228 274 BY SIMILARITY.
FT DISULFID 273 281 BY SIMILARITY.
FT DISULFID 293 307 BY SIMILARITY.
FT DISULFID 306 317 BY SIMILARITY.
FT DISULFID 344 389 BY SIMILARITY.
FT DISULFID 388 397 BY SIMILARITY.
FT DISULFID 420 466 BY SIMILARITY.
FT DISULFID 465 476 BY SIMILARITY.
FT DISULFID 489 505 BY SIMILARITY.
FT DISULFID 504 515 BY SIMILARITY.
FT DISULFID 542 587 BY SIMILARITY.
FT DISULFID 586 595 BY SIMILARITY.
FT CARBOHYD 500 500 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLUCT 24 24 F -> M (IN REF. 3).
SQ SEQUENCE 615 AA; 69918 MW; E59E4BBAEC066C6 CRC64;
Query Match 50.28; Score 1557.5; DB 1; Length 615;
Best Local Similarity 46.74; Pred No. 4.3e-94;
Matches 273; Conservative 118; Mismatches 192; Indels 1; Gaps 1;
3 HKSEVAHRFDLGENFKALVLIYAQYLOCCPPEDHVKNVETFEPAKTCVADESAENC 62
30 HKSEIARVNDLKEETKAVAMI:FAQYLQRCSEGLSKLVKQWDLAQKCVANEDAPEK 89
63 DKSHTLPGDKLTVATLRETYGEMACCAKQPERNECFCHKDDPNLPR-LVPRFVD 121
90 SKPLPSILIOEQVEXLRDSYGAVACCSKADPERNECFLSFKVSPQDFVQVQPEASD 149
122 VMCTAFHNEETFLKVLVLELARHPVFYAFELLFFPAKVKYKAFTCCQAADKAACLPK 161
150 VICEYQDNVSTFGHTFYSVARRHPFLYAPAILSTFVDFEHALQSCCKSDVGACLDTK 209
182 LDELDEKASAKQRLKASLQKFGERAFAKAVARLSQKFPKAEFAVSK-VTDLTKV 241
210 EIVMREKAGSVYKQVFCGLIKQGDVRFQARQLIV:SQYKPAFSEVSKFVHDSIGV 269
242 HTECHGDLLECADRADRLAKYICENQDS:SSKKECECEKPLEKSHCIABVENDSMFAD 301
270 HKECEGDMVECYEDMARMMSNCSQCVFSGKIKDCEKPIVERSQCIINEAEFDEXPAD 329
302 LPSLAADVEESKQVKNVAEAKQVFLGNFLVEYARRHPDYSVLLBLAKTYETTLKCC 361
330 LPSLVKYEIECKVECKSPFAGHDAPFAEFVVEYRRHPFES:QLIMRIAGYESLLEKCC 389

```

```

QY 362 AAADPHCYAKVPBEPKPLVEEPQNLIKONCELFQEQJSEYKFNALLVRYTKKVPQVSTP 421
DB 390 KTDNPAECYANAQOLNQH:KETQDVVKNTCDLLHDH3EADFLKSLIRY:KQMPQVPTD 449
QY 422 TLVEVSRNLGVSKCKCKHPEAKYPCAEADY:SVVLNQLCVLHEKTPVSDVTKCTTESL 481
DB 450 LLETGKXMTTIGTKCCQLGDERRMACSEGLSVIHDTCRKQETTPINDKVSQCCSOLY 509
QY 482 VNRRPCSALEVDYTPVKEFNAETTFHADICTLSEKERICKQTALVELVVKHKPKATK 541
DB 510 ANRRPCFTAMGVDTKYVPPPNPMFSEDEXLCSAPAEEREVGOMKLLNLIKRPQWTE 569
QY 542 EQLKAVMDOFAAFVEKCKKADKCTCFABEOKKJVAASQAALGL 585
DB 570 EQIKTIADGFTAMVDKCKCKQSDINTCFEGGANLIVQSRATLGI 613

RESULT 14
FETA PANTR
ID FETA PANTR STANDARD; PRT: 609 AA.
AC Q28789;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-
DE fetoprotein).
GN AFP.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96032345; PubMed=7557431;
RA Nisio H., Gibbs P.E., Minghetti P.P., Zielinski R., Dugaiczky A.;
RT "The chimpanzee alpha-fetoprotein-encoding gene shows structural
RT similarity to that of gorilla but distinct differences from that of
RT human.";
RL Gene 162:213-220(1995).
CC -!- FUNCTION: BINDS COPPER, NICKEL, AND FATTY ACIDS AS WELL AS, AND
CC BILIRUBIN LESS WELL THAN, SERUM ALBUMIN. ONLY A SMALL PERCENTAGE
CC (LESS THAN 2%) OF THE HUMAN AFP SHOWS ESTROGEN-BINDING PROPERTIES.
CC -!- SUBUNIT: DIMERIC AND TRIMERIC FORMS HAVE BEEN FOUND IN ADDITION
CC TO THE MONOMERIC FORM.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and
CC yolk sac.
CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U21916; AAA91641.1; -
CC PIR: JC4258; JC4258.
CC HSP: P02768; IE7B.
CC InterPro: IPR000264; Serum_albumin.
CC Pfam: PF00273; transport_prot; 3.
CC PRINTS: P008002; SERUMALBUMIN.
CC ProDom: PD002486; Serum_albumin; 1.
CC SMART: SM00103; ALBUMIN; 3.
CC PROSITE: PS00212; ALBUMIN; 2.
CC Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;
KW Signal.
FT SIGNAL 1 18 BY SIMILARITY.
FT CHAIN 19 609 ALPHA-FETOPROTEIN.
FT DOMAIN 20 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.

```

FT	DOMAIN	404	595	ALBUMIN 3.
FT	METAL	22	22	COPPER AND NICKEL (BY SIMILARITY).
FT	DISULFID	99	114	BY SIMILARITY.
FT	DISULFID	133	124	BY SIMILARITY.
FT	DISULFID	148	193	BY SIMILARITY.
FT	DISULFID	192	201	BY SIMILARITY.
FT	DISULFID	224	270	BY SIMILARITY.
FT	DISULFID	269	277	BY SIMILARITY.
FT	DISULFID	289	303	BY SIMILARITY.
FT	DISULFID	302	313	BY SIMILARITY.
FT	DISULFID	384	393	BY SIMILARITY.
FT	DISULFID	416	462	BY SIMILARITY.
FT	DISULFID	461	472	BY SIMILARITY.
FT	DISULFID	485	501	BY SIMILARITY.
FT	DISULFID	500	511	BY SIMILARITY.
FT	DISULFID	538	583	BY SIMILARITY.
FT	DISULFID	582	591	BY SIMILARITY.
FT	CARBOHYD	42	42	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	251	251	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SEQUENCE	603 AA;	68741 MW;	C032987:CAD92672B CR664;

  

Query Match	40.4%;	Score 1253.5;	DB 1;	Length 609;
Best Local Similarity	40.1%;	Pred. No. 2.4e-74;		
Matches 236;	Conservative 115;	Mismatches 231;	Indels 7;	Gaps 30;

  

QY	3	HASE-----VAHRFKDLGEENFKALVLIARAOYLQCCPFEDHVKLVNEVTEFAKCVADE	57
DB	22	HRNEVGYASILDSYCTAEINLTDLAT:FAQVQGEATYKEVSKVKVDALTA:EKPTGDE	81
QY	58	SAENCDKSLHTLFGKLCVTATLETETGYEMADCCAKQOEPERNECFLOHKDKNP-NLPLV	116
DB	82	QSAGCLENGLPAFL-EELCREKEILUEKYGH-SDCSOSSEGRHNCPLAHKKPTPASPFPFQ	140
QY	117	REPVCMCTAFHCBETLKKLYLEIARRHPYFVARELLFPAPKRYKAAFTESCQAADKAA	176
DB	141	VBPVTSGEAEEDEETPMNKPFIYEARRHPFLYAPTILILWAARYDK:IPSCCKRAENAVE	200
QY	177	CLPLKMLDELRDEGKASSAKQRLKCAKCKTGERAFKAMAVARLSQRPFXAEFAEYKLVV	236
DB	201	CFCTKAAVTYKRELRESSLNLQHACAVMKNEGTRTFQAITVTKLSQFTKVNFTETQKLV	260
QY	237	DLTKVHTCCCHDLECCADDRADIAKYICENQDSISKKKECCERKPLLEKSHCIAEVND	296
DB	261	DVARVHEHCRCRGVLDCLQDGEKIMSY:CSQDQTLNKKITCCCKLTLESGCIIHAEND	320
QY	297	ENPADLPSLAADFVESKVCVKYAEADVFLGMFLVEYARRHPDYSVW:LJLAKTYET	356
DB	321	EXPEGLSPNLNEFLGCRDFNPFSSGEKNIFLASVHEYSRRHPQLAVSVILAVAKGYQL	380
QY	357	LEKCCAAADPHCYAKVDEFKPLVEPQNLIKONCELPQCLGEYKFNALLVRYTKXVP	416
DB	381	LEKCFOTENPLECQDGBEELQKYIQESQALAKRSGLFQKLGEYLYONAFVAYTKAP	440
QY	417	QVSTPTLVEVSNLKVGSCKCKHPEAKRMPCEADYLSVVLNOLCVLHEKTPVSRVTKC	476
DB	441	QLTSSSELMAITRKMAATAATCCQLSEDKLACGEGAADIIGHLCIRHETTPNPGVQC	500
QY	477	CTESLVNRRPCFSALEVDETYVPKFEFNAETFTPEADICTLSEKEROIKKQATALVELVKHK	536
DB	501	CTSSYANRRPCFSSLVWDETVPVPAFSDDKFIHKDLCCQAQVALQTMQOEFLINLVKQK	560
QY	537	PXATKEQLKAVMDDFAAFVEKCCKADDKETCFABEGKLVAAASQAALGL	595
DB	561	PQITSEQLAEVIAIDFSGLEKCCQGEQEVCFABEGOKLISKTRAAALGV	609

  

RESULT 15	
FETA_HUMAN	
ID_FETA_HUMAN	STANDARD;
AC	P02771;
DT	21-JUL-1986 (Rel. 01, Created:
DT	21-JUL-1986 (Rel. 01, Last sequence update:
DT	28-FEB-2003 (Rel. 41, Last annotation update)

RT flanking region.";  
 RL J. Biol. Chem. 283:5055-5060(1985).  
 RN [10].  
 RP METAL-BINDING.  
 RX MEDLINE=79001617; PubMed=80265;  
 RA Aoyagi Y., Ikenaka T., Ichida F.;  
 RT "Copper(II)-binding ability of human alpha-fetoprotein.";  
 RL Cancer Res. 38:3483-3486(1978).  
 RN [11].  
 RP BILIRUBIN-BINDING.  
 RX MEDLINE=80001710; PubMed=89902;  
 RA Aoyagi Y., Ikenaka T., Ichida F.;  
 RT "Alpha-Fetoprotein as a carrier protein in plasma and its bilirubin-binding ability.";  
 RL Cancer Res. 39:3571-3574(1979).  
 RN [12].  
 RP SULFATION.  
 RX MEDLINE=86042625; PubMed=2414772;  
 RA Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;  
 RT "Tyrosine sulfation of proteins from the human hepatoma cell line HepG2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).  
 CC -!- FUNCTION: BINDS COPPER, NICKEL, AND FATTY ACIDS AS WELL AS, AND BILIRUBIN LESS WELL THAN, SERUM ALBUMIN. ONLY A SMALL PERCENTAGE (LESS THAN 2%) OF THE HUMAN AFP SHOWS ESTROGEN-BINDING PROPERTIES.  
 CC -!- SUBUNIT: DIMERIC AND TRIMERIC FORMS HAVE BEEN FOUND IN ADDITION TO THE MONOMERIC FORM.  
 CC -!- TISSUE SPECIFICITY: Secreted.  
 CC -!- TISSUE SPECIFICITY: PLASMA. SYNTHESIZED BY THE FETAL LIVER AND YOLK SAC.  
 CC -!- DEVELOPMENTAL STAGE: OCCURS IN THE PLASMA OF FETUSES MORE THAN 4 WEEKS OLD, REACHES THE HIGHEST LEVELS DURING THE 12TH-16TH WEEK OF GESTATION, AND DROPS TO TRACE AMOUNTS AFTER BIRTH. THE SERUM LEVEL IN ADULTS IS USUALLY LESS THAN 40 NG/ML. AFP OCCURS ALSO AT HIGH LEVELS IN THE PLASMA AND ASCITIC FLUID OF ADULTS WITH HEPATOMA.  
 CC -!- PTM: INDEPENDENT STUDIES SUGGEST HETEROGENEITY OF THE AMINO-TERMINAL SEQUENCE OF THE MATURE PROTEIN AND OF THE CLEAVAGE SITE OF THE SIGNAL SEQUENCE.  
 CC -!- PTM: SULFATED.  
 CC -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.  
 CC -!- SIMILARITY: Contains 3 albumin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M10949; AA451674.1; -  
 DR EMBL; M10950; AA451675.1; -  
 DR EMBL; V21514; CA24756.1; -  
 DR EMBL; M16110; AA38754.1; -  
 DR EMBL; Z19532; CA479592.1; -  
 DR PIR; A26524; FPHJ.  
 DR HSSP; PC2768; 1E7S.  
 DR GlycoSuiteDB; P02771; -  
 DR Siena-2DPAGE; P02771; -  
 DR Genew; HGNC:317; AFP.  
 DR MIM; 104150; -  
 DR InterPro; IPR000264; Serum\_albumin.  
 DR Pfam; PF00273; transport\_prot; 3.  
 DR PRINTS; PR08002; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum\_albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 2.  
 DR Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;  
 KW Signal; Polymorphism.  
 FT SIGNAL; 18  
 FT CHAIN; 19 609 ALPHA-FETOPROTEIN.  
 FT DOMAIN; 20 205 ALBUMIN 1.  
 FT DOMAIN; 212 397 ALBUMIN 2.

FT	DOMAIN	404	595	ALBUMIN 3. COPPER AND NICKEL.
FT	METAL	22	114	
FT	D-SULFID	99	124	
FT	D-SULFID	113	193	
FT	D-SULFID	148	201	
FT	D-SULFID	224	277	
FT	D-SULFID	289	303	
FT	D-SULFID	302	313	
FT	D-SULFID	384	393	
FT	D-SULFID	416	462	
FT	D-SULFID	461	472	
FT	D-SULFID	485	501	
FT	D-SULFID	500	511	
FT	D-SULFID	538	583	
FT	D-SULFID	582	591	
FT	CARBOHYD	251	251	
FT	VARIANT	570	570	
FT	SEQUENCE	609 AA;	68677 MW;	404E45820E1C2D4F CRC64;
SC	SEQUENCE	609 AA;	68677 MW;	404E45820E1C2D4F CRC64;
Query Match		40.3%;	Score 1249.5;	DB 1; Length 609;
Best Local Similarity		39.9%;	Pred. No. 4.4e-74;	
Matches		235;	Conservative 116;	Mismatches 231; Indels 7; Gaps 3;
Qy	3 HKSE-----VAHRFKDLGEENFKALVLIATAFYLOQCPFDHVKLVNVEYFEAKTCVADE	57		
Db	22 HRNEYGASILDSYQCTAEISLADLATIFFAQVQVQATYKEVSKMVKDALTAIEKPTGDE	81		
Qy	58 SARNDKSLHTJFGDKLCTVATLRETYGEMADCCAXQEPERNECFLOHXDNP-NILPLV	116		
Db	82 QSSGCLNQLPAFLBLCHEKEILEKYGH-SDDCSQSEEGHNCFLAKKKPTPASIPLEQ	140		
Qy	117 RPEVDVMTAFHONBETFLKKYLYEARRHPYAPYAPELLFFAKRYKAAFECCQAADKAA	176		
Db	141 VPBFTVSCAEYEDRETFMKNKFIYARRHPFLYAPTILLMAARYDKIIPSCCKAENAVE	200		
Qy	177 CLLPKLDELDEGKASSAKQRLCASLQKFGERAKAVARLSQRPFAEFAEVSGLVT	236		
Db	201 CFQTKATVTKEIRRESLLNQHACAVMKNGTTFQAITVKLSQKFTKYNFTIOLVL	260		
Qy	237 DLTKVHTECHGDLLECCADRDALAKYICENQDSISSKKECEKPLEKSHCIAEVND	296		
Db	261 DVAHVHEHCCRGDVLDCLODGEKIMSYCSQDTLSNKTCECKLTJERGQCIIHAEND	320		
Qy	297 EMPADLPSLAADFESKDVCKNYABAKDVLGMLFELYARHHPDYVVLJLRKATYETT	356		
Db	321 EKPEGLSPNLRFLGDRDFNQFSSGKNITFLAGFVHEYSRRHPQLAVSVTLRVAKGQEL	380		
Qy	357 LEKCCAAAOPHECYAKVDFEFKPLVEEPQNLIKONCELFELQGEYKFNQALJLVRYTKVP	416		
Db	381 LEKCFOTENPLECQDKEELQKYIGESQALAKSSCGLFQKLGYYLQNAFJVAITKAP	442		
Qy	417 QVSTPTLVEVSRNLGVGSKCKHPKAPKRYPCAEYLSVVLNQLCVLHEKTPVSRVTKC	476		
Db	441 QLTSSSELVAITRKMAATATATCCQLSEDKLACGEGAADIIGHLCIRHEMTFVNPVGQC	500		
Qy	477 CTESLVNRRRCFSALEVDITYVFKENAEFTFHADICITLSEKERQIKKTALVELVKHK	536		
Db	501 CTSSYANRRRCFSSLVVDITYVPAPFSDDKFIHFHKLCOAQGVVALQTMQOEFLINLVKQK	560		
Qy	537 PKATKEOLKAVMDDFAAFEVKECKCKADKFCFAEEGKGLVAASQAALGL	585		
Db	561 PQTTEELEAVIADFSGLLEKCKCQGEQVEVCFABEGCKLISKTRALGV	609		

Search completed: October 27, 2003, 15:19:12  
 Job time : 27 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein: - protein search, using sw model

Run on: October 27, 2003, 15:15:23 ; Search time 106 Seconds  
(without alignments)  
1424.156 Million cell updates/sec

Title: US-09-833-117-18

Perfect score: 3103

Sequence: 1 DAHSEVAHRFKDLGLENFK.....TCFAEKGKLVAAQAALG: 585

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SF:REMBL\_23.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_prodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2504	80.7	608	5 Q95VB7	Q95vb7 schistosoma
2	2374	76.5	608	11 C8C7H3	C8c7h3 mus musculus
3	2336	75.3	576	1: Q8C7C7	Q8c7c7 mus musculus
4	1865.5	60.1	396	4 Q8IUK7	Q8iuk7 homo sapien
5	1295.5	41.7	527	13 Q8JIA9	Q8jia9 sphendon p
6	1242	40.0	609	6 Q8MJU5	Q8mju5 canis famil
7	1242	40.0	626	13 Q8JW05	Q8jw05 ambystoma m
8	1218.5	39.3	610	6 Q8MJ76	Q8mj76 sus scrofa
9	1087	35.0	624	13 C8JW06	Q8jw06 ambystoma t
10	1082	34.9	605	11 C8BK65	C8bk65 mus musculus
11	1076	34.7	605	11 C8BK56	C8bk56 mus musculus
12	1045	33.7	400	13 Q8JIA7	Q8jia7 sphendon p
13	955	30.8	603	13 Q9YGH6	Q9ygh6 rana shqipe
14	928.5	29.9	614	13 Q91134	Q91134 naja naja
15	888	28.6	406	13 Q8JIA8	Q8jia8 hepiodactyl
16	739	23.8	235	11 Q8CG74	Q8cg74 mus musculus

17	713	23.0	47.7	11 Q8R0J9	Q8r0j9 mus musculu
18	376.5	12.1	484	13 Q9W6F5	Q9wf65 gallus gall
19	373	12.0	476	11 Q9CY31	Q9cy31 mus musculu
20	372	12.0	476	11 Q91XG1	Q91xg1 mus musculu
21	365	11.8	551	13 Q42279	Q42279 petromysc
22	331	10.7	122	13 Q90WZ8	Q90wz8 iarus argen
23	290	9.3	123	13 Q90WZ6	Q90wz6 poephila gu
24	264	8.5	135	11 Q63205	Q63205 rattus ncrv
25	186	6.1	1723	2 Q9JMX8	Q9jmx8 helicobacte
26	184	5.9	1819	16 Q9ZLV0	Q9zlv0 helicobacte
27	184	5.9	1927	16 Q25262	Q25262 helicobacte
28	162.5	5.2	44	6 Q95MC2	Q95mc2 equus cabal
29	162.5	5.2	680	5 Q9V6S8	Q9v6s8 drosophila
30	161	5.2	1079	3 Q96V11	Q96v11 pneumocysti
31	156	5.0	1026	3 Q74669	Q74669 pneumocysti
32	154.5	5.0	3843	5 Q9USD0	Q9usd0 drosophila
33	153.5	4.9	661	5 Q8MS79	Q8ms79 drosophila
34	153.5	4.9	3843	5 Q9VU94	Q9vu94 drosophila
35	153	4.9	62	6 Q8MIL1	Q8mil1 sus scrofa
36	150	4.8	1065	3 Q01828	Q01828 pneumocysti
37	149	4.8	40	6 Q9TRA5	Q9tra5 cryptoslagus
38	148	4.8	1028	3 Q74668	Q74668 pneumocysti
39	144.5	4.7	8749	4 Q8NF91	Q8nf91 homo sapien
40	142.5	4.6	2756	10 Q9LJ60	Q9lj60 arabidopsis
41	141.5	4.6	1560	5 Q26644	Q26644 strongyloce
42	138.5	4.5	1069	3 Q96V12	Q96v12 pneumocysti
43	137.5	4.4	1348	16 Q8YK55	Q8yk55 anabaena sp
44	135.5	4.4	2841	5 Q8MLU9	Q8mlu9 drosophila
45	135.5	4.4	2931	5 Q9W2C6	Q9w2c6 drosophila

#### ALIGNMENTS

RESULT 1

Q95VB7 Q95VB7 PRELIMINARY; PR: 608 AA.

AC Q95VB7; AC Q95VB7; 19, Created;

DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update);

DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update);

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update);

DE Albumin.

OS Schistosoma mansoni (Blood fluke).

OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;

OC Schistosomatidae; Schistosomatidae; Schistosoma.

OX NCBI\_TaxID=6483;

RN [1]\_TaxID=6483;

RP SEQUENCE FROM N.A.

RA Osmar A., Asahi H., Stadecker M.J., Loverde P.T.;

RT "Albumin precursor homolog is a novel T helper cell immunogenic egg

PT component in murine infection with Schistosoma mansoni.";

RL Submitted (SEP-2001) to the EMBL/GenBank/DBSJ databases.

DR EMBL; AF418550; AAL08579.1; ..

DR InterPro; IPR000264; Serum\_albumin.

DR Pfam; PF00273; transport\_prot; 3.

DR PRINTS; PR00802; SERUMALBUMIN.

DR ProDom; PDC02486; Serum\_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PSC0212; ALBUMIN; 2.

SQ SEQUENCE 608 AA; 68225 MW; E5EABB28E1C66E54 CRC64;

Query Match 80.7%; Score 2504; DB 5; Length 608;  
Best Local Similarity 76.3%; Pred. No. 1.7e-188;  
Matches 445; Conservative 79; Mismatches 59; Indels 0; Gaps 0;

OY 1 DAHSEVAHRFKDLGLENFKLVIAFAVYLCQCFEDHVKLVNEVTEFAKTCVADESAAE 60

DB 25 DAHSEIAHREKDLGEQHPKGLVLIASFQELQKCPYEHVKLVNEVTEFAKTCVADESAAE 84

OY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKODNPNLPLVRPEV 120

DB 85 NCDKSLHTLFGDKLCAIPTLRLDSYGLADCCAKKEPRNECFLOHKODHNPPLPFFVRPDA 144

```

QY 121 DVMCTAFHDNEETFLKXYLYEIAARRHPYFYAPPELLFFAKDYKAAFTBCCCAQAAKACLLP 180
DB 145 EAMCTSFQENAVTFMGHYLHEVARRRHPYFYAPPELLFFAAYAEKYSAINFTBCCGEADKAACITP 254
QY 181 KJDELREDEKASSAKORLKASLOKGERAFKAWAVARLSQTFPNACPAEITKLATDUTK 240
DB 205 KJDAKBAKALASSVNGKULSSQRFQORAFKAWAVARMSQKFPKADFAEITKLATDUTK 254
QY 241 VHTBECCHGDLLECADRADLAKYICENQDSISSKJAECCCKPLLEKSHCIAEVENDEMPA 300
DB 265 LTBECCHGDLLECADRADLAKYICENQASISSKLOACCKEVLKXSHCLSEVENDELPA 324
QY 301 DLPSLAADFVESKDYCKNYAEAKOVFLGMYEYARRHPDYVSVLLRLAKTYETTLKXC 360
DB 325 DLPSLAADFVEDCKVCYKNAEAKOVFLGTLFYEYARRHPDYVSVALLRLAKTYEATLEK 364
QY 361 CAADPHCEYAKVDFEFPKPLVEBPONLJKONCELFEOLGEVKEKQNALVRYTKVPQVST 420
DB 385 CAEADPSACYGKVJDEFQPLVEBPONLVKANCELFELKLGEGYGFONALVRYTKVQAPVST 444
QY 421 PTJVEVSRNLGKVGSKCKGPEAKEMPCAEDYLSWJNQLCVJHEKTPVSDRVTKCTES 480
DB 445 PTJVEAARNLGKVGSKCVLPKAEGLSCQVEDYISAILNRVCVJHEKTPVSEQVTKCTGS 504
QY 481 LVNRRPCFSALEVDETVVPKEFNASTFTPHADICTLSEKERQIKKQATALVELVKHKPKAT 540
DB 505 VERRPCFSALPVJDETVVPKEFKAETTFHSDICTLPEKEKQIKKQATALVELVKHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADCKETCPABEGKKLVAASQAAL 583
DB 565 GPQJRTVLGEFTAFDLKCKKAEADKAEFSDGPKLVASSQAAL 607

RESULT 2
Q8C7H3 PRELIMINARY; PRT; 608 AA.
ID Q8C7H3 AC Q8C7H3:
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Albumin 1 (Fragment);
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
DR Nature 420:563-573(2002);
DR EMBL; AK050248; BAC34145.1; -.
SQ SEQUENCE 608 AA; 68722 MW; 292F603BED3A61B4 CRC64;

Query Match 76.5%; Score 2374; DB 11; Length 608;
Best Local Similarity 72.2%; Pred. No. 2.7e-178;
Matches 421; Conservative 80; Mismatches 82; Indels 0; Gaps 0;

QY 1 DAHKEVARRFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADSESAE 60
DB 25 EAHKSEIAHRYNDGEQHFEGKGLVLIAPFQYLQKCSYDEHAKLVQEVTDFAKTCVADSESA 84
QY 61 NCKDSLHTLFGDKLCTVATLSETYGMADCCAKQEPERNECFLOHKQDNPNLPLRVPEV 120
DB 85 NCKDSLHTLFGDKLCAINLRENYGELADCTCKQEPERNECFLOHKQDNPNLPLRVPEV 144
QY 121 DVMCTAFHDNEETFLKXYLYEIAARRHPYFYAPPELLFFAKRYKAAFTBCCCAQAAKACLLP 180
DB 145 EAMCTSFQENAVTFMGHYLHEVARRRHPYFYAPPELLFFAAYAEKYSAINFTBCCGEADKA 204

```

```

QY 181 KJDELREDEKASSAKORLKASLOKGERAFKAWAVARLSQTFPNACPAEITKLATDUTK 240
DB 205 KJDAKBAKALASSVNGKULSSQRFQORAFKAWAVARMSQKFPKADFAEITKLATDUTK 254
QY 241 VHTBECCHGDLLECADRADLAKYICENQDSISSKJAECCCKPLLEKSHCIAEVENDEMPA 300
DB 265 LTBECCHGDLLECADRADLAKYICENQATSSKLOACCKEVLKXSHCLSEVENDELPA 324
QY 301 DLPSLAADFVESKDYCKNYAEAKOVFLGMYEYARRHPDYVSVLLRLAKTYETTLKXC 360
DB 325 DLPSLAADFVEDCKVCYKNAEAKOVFLGTLFYEYARRHPDYVSVALLRLAKTYEATLEK 364
QY 361 CAADPHCEYAKVDFEFPKPLVEBPONLJKONCELFEOLGEVKEKQNALVRYTKVPQVST 420
DB 385 CAEANPPACYGKVLAEFQPLVEEPKLVKTNCDLYEKLGEYGFONALVRYTKVQAPVST 444
QY 421 PTJVEVSRNLGKVGSKCKGPEAKEMPCAEDYLSWJNQLCVJHEKTPVSDRVTKCTES 480
DB 445 PTJVEAARNLGKVGSKCCTLPEDQKPCVEDYLSAILARVCLLREKTPVSEVTKCCSGS 504
QY 481 LVNRRPCFSALEVDETVVPKEFNASTFTPHADICTLSEKERQIKKQATALVELVKHKPKAT 540
DB 505 LVERRPCFSALTVDETVVPKEFKAETTFHSDICTLPEKEKQIKKQATALVELVKHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADCKETCPABEGKKLVAASQAAL 583
DB 565 ABQLKTVMDFAQFDTCCAKAADCKOTCFSTEGPNLVTRCKDTL 607

RESULT 3
Q8C7C7 PRELIMINARY; PRT; 576 AA.
ID Q8C7C7 AC Q8C7C7:
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Albumin 1 (Fragment);
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
DR Nature 420:563-573(2002);
DR EMBL; AK050644; BAC34360.1; -.
FT NON_TER 1
SQ SEQUENCE 576 AA; 65002 MW; F85733E99AE37F04 CRC64;

Query Match 75.3%; Score 2336; DB 11; Length 576;
Best Local Similarity 72.2%; Pred. No. 2.5e-175;
Matches 415; Conservative 79; Mismatches 81; Indels 0; Gaps 0;

QY 9 HRFKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADSESAENCCKSJHT 68
DB 1 NRYNDLGEQHFEGKGLVLIAPFQYLQKCSYDEHAKLVQEVTDFAKTCVADSESAENCCKSJHT 60
QY 69 LFGDKLCTVATLSETYGMADCCAKQEPERNECFLOHKQDNPNLPLRVPEVVMCTAFH 128
DB 61 LFGDKLCAINLRENYGELADCTCKQEPERNECFLOHKQDNPNLPLRVPEVPEAEAMCTSEK 120
QY 129 DREETFLLKXYLYEIAARRHPYFYAPPELLFFAKRYKAAFTBCCCAQAAKACLLPKLDELRE 188
DB 121 ENPTFMGHYLHEVARRRHPYFYAPPELLFFAAYAEKYSAINFTBCCGEADKAEKSCUTPK 180
QY 189 GXASSAKORLKASLOKGERAFKAWAVARLSQTFPNACPAEITKLATDUTKVNKECCHG 248
DB 181 ALVSSVRRQKSSMOKFGERAFKAWAVARLSQTFPNACPAEITKLATDUTKVNKECCHG 240

```

QY 249 DLLECAADDRADJAKY-CENQDSISSKLECCCEKPLLEKSHCIAEVENDEMPADLSLAAD 308  
 DB 241 DLLECAADDRADJAKY-CENQAT-SSKQTCDDPLKKAHCHSEVBDHTMPADLSLAAD 300  
 QY 309 FVSEKDVCKNYAEAKDVLGMFLYIYARRHPDYVSULLRLAKTYETTT-EKCCAAADPHE 368  
 DB 302 FVEDQEVCKNYAEAKDVLGTLFLYIYARRHPDYVSULLRLAKTYEATLEKCCAAEAPPA 360  
 QY 369 CYAKVDFEPKPLVSEPNLIKONCELFQELGEYKFNALLVRYTKKVPQVSTPTLVEVSR 428  
 DB 361 CYGTVLAEFQPLVEEPKPLVKTCNCLYELKGLGEYGFQNAIVRYTKAPQVSTPTLVEAAR 420  
 QY 429 NLGKVGSKCKHPEAKRMPCAEADYLSVNLQCLVLEKTPVSDRVTKCCCTESLVNRRPCF 488  
 DB 421 NLGRVGTCKCTLPEDGRPLCPVEDYLSAILNRVCLLHEKTPVSHVTKCCSGSLVRRPCF 480  
 QY 489 SALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVHKHPKATKEQLKAYM 548  
 DB 481 SAUTVDETYVPKEFNAETFTFHSDICTLPEKEQIKKQATALVELVHKHPKATAEQLKAYM 543  
 QY 549 DPEAAAEVEKCKKADDEKTCFAEEGKKLVAASQAAL 583  
 DB 541 DPEAQFLDTCCAAAGKOTCFSTEGPNLVTRCKDAL 575

## RESULT 4

Q81UK7 PRELIMINARY; PRT; 396 AA.  
 ID Q81UK7  
 AC Q81UK7  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Similar to serum albumin precursor.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Strausberg R.;  
 RJ Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC035969; AAH35969.1;  
 SQ SEQUENCE 396 AA; 45153 MW; 756519C096463A9B CRC64;

Query Match 60.1%; Score 1865.5; DB 4; Length 396;  
 Best Local Similarity 63.6%; Pred. No. 1.6e-138;  
 Matches 372; Conservative 0; Mismatches 0; Indels 213; Gaps 1;

QY 1 DAKSEVAHRFKDLGRENFKALVLIAPQVLOQCPFDHVKLVNEVTEFAKTCVADESSE 60  
 DB 25 DAKSEVAHRFKDLGRENFKALVLIAPQVLOQCPFDHVKLVNEVTEFAKTCVADESSE 84  
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 120  
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 144  
 QY 121 DVMTAFHDNEETFLKXYLIEARRHPDYFAPPELLFFAKRYKAAFTCCCAADKAACLLP 180  
 DB 145 DVMTAFHDNEETFLKXYL  
 QY 181 KLDELROEGKASSAKORLKASLQKFGERAFKAWAVARLSQRPFAEFAVSKLVDTLTK 240  
 DB 164 ----- 163  
 QY 241 VHTCCCHGDLLECCADDRADJAKYICENQDSISSKLECCCEKPLLEKSHCIAEVENDEMPA 300  
 DB 164 ----- 163  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYVSULLRLAKTYETTTLEK 360  
 DB 164 ----- 163  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYVSULLRLAKTYETTTLEK 360  
 DB 164 ----- 163  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYVSULLRLAKTYETTTLEK 360  
 DB 164 ----- 163

QY 361 CAAADPHECYAKVDEPKPLVSEPNLIKONCELFQELGEYKFNALLVRYTKKVPQVST 420  
 DB 172 CAAADPHECYAKVDEPKPLVSEPNLIKONCELFQELGEYKFNALLVRYTKKVPQVST 231  
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVNLQCLVLEKTPVSDRVTKCCCTES 480  
 DB 232 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVNLQCLVLEKTPVSDRVTKCCCTES 291  
 QY 481 LVNRRPFCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVHKHPKAT 540  
 DB 292 LVNRRPFCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVHKHPKAT 351  
 QY 541 KEQLKAYMDPEAAAEVEKCKKADDEKTCFAEEGKKLVAASQAALGL 585  
 DB 352 KEQLKAYMDPEAAAEVEKCKKADDEKTCFAEEGKKLVAASQAALGL 396

## RESULT 5

Q8JIA9 PRELIMINARY; PRT; 527 AA.  
 ID Q8JIA9  
 AC Q8JIA9  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Serum albumin (Fragment).  
 OS Sphenodon punctatus (Hastaria) (Tuatara).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.  
 OX NCBI\_TaxID=8508;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Vercalis V.J., Brennan S.O., George P.W., Chambers G.K.;  
 RT "Partial coding sequence for Sphenodon punctatus 68 kDa albumin."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF375971; AAM46104.1;  
 DR IPI:IPRO; IPR000264; Serum albumin.  
 DR Pfam; PF00273; transport prot; 3.  
 DR PRINTS; PR03802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum\_albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 3.  
 FT NON TER 1  
 SQ SEQUENCE 527 AA; 59711 MW; C62B799E387F5929 CRC64;

Query Match 41.7%; Score 1295.5; DB 13; Length 527;  
 Best Local Similarity 45.9%; Pred. No. 1.4e-93;  
 Matches 240; Conservative 92; Mismatches 198; Indels 3; Gaps 3;

QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVREPV 120  
 DB 5 CJKSLDTIFLDEICHEEGFAAKY-DLAACCAKAEVERKECLLAHKNATPGFAPQRPPI 63  
 QY 121 DVMTAFHDNEETFLKXYLIEARRHPDYFAPPELLFFAKRYKAAFTCCCAADKAACLLP 180  
 DB 64 EVSCCKLQYDDRLTLGNGYIYEVARRHPYLVPPVFATASLYDEALKTCCTADKATCFHP 123  
 QY 181 KLDELROEGKASSAKORLKASLQKFGERAFKAWAVARLSQRPFAEFAVSKLVDTLTK 240  
 DB 124 RIPLLEIYLVKMSNGIOENTCGILKKFGERTLAKTLQVMSOKPFPADFATINKLVEDITH 163  
 QY 241 VHTCCCHGDLLECCADDRADJAKYICENQDSISSKLECCCEKPLLEKSHCIAEVENDEMPA 300  
 DB 184 MHTCCCHGDLLECCADDRADJAKYICENQDSISSKLECCCEKPLLEKSHCIAEVENDEMPA 243  
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYIYARRHPDYVSULLRLAKTYETTTLEK 360  
 DB 244 DLSERIAEYEDPHVCDHLAKEQDAFLAKFLYIYARRHPDYVSULLRLAKTYETTTLEK 303  
 QY 361 CAAADPHECYAKVDEPKPLVSEPNLIKONCELFQELGEYKFNALLVRYTKKVPQVST 420  
 DB 304 CKTDNPPCYQAAADLKHGIAQFQELVQNCDDYNTLGGYLFHALLIRYTKRMPGLTS 363



```

QY 355 TTLEKCCAAADPHECYAKVDFEKPPLVEEPONLIKONCELFEGLEYKFKONALLHRYTKK 414
Db 388 EFLKCCAAAGHNECCAKTESKKEIESSVTJKTNCAGALCKGKSLFONLLIFKIVAR 447
QY 415 VPOVSTPTLVEVRNRLKGVSKCKHPKAPKMPKADYLSWLNQCLVLHEKTPVSDRYT 474
Db 448 MPALSEGSLRLITKSMITIGEKCHRPEDQOMTCSEGLGVFGOICMKCKTTPUNEKVA 507
QY 475 KCTESLNNRRPFCFSALEVEDTVPKEFNAETTFHADICTLSEKROKKQCTALVEJYK 534
Db 508 QCCSHSLSSQTPCFSAIPVDETVVPPPLSVASFNFNDELCTTSEPQQSKQVFLIRLMK 567
QY 535 HKPKATKEQLKAVMDPFAEVKCCXADDKETCFABEGKLVAAASOALGL 585
Db 568 QYRMTDEQLKTCVNVFVPMVQCCXADNHNCFALGAKLIDACKAILAV 618

RESULT 8
Q8MJ76 PRELIMINARY; PRT; 610 AA.
AC Q8MJ76
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Alpha-fetoprotein.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim J.G., Noncena D., Vailet J.L., Christenson R.K.:
RT "Mapping of the porcine alpha-fetoprotein (AFP) gene to SSC8.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF517776; AAM56110.1; -.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 610 AA; 68624 MW; C985BEAD44963D5E CRC64;

Query Match: 39.3%; Score 1218.5; DE 6; Length 610;
Best Local Similarity 39.7%; Pred. No. 19e-87;
Matches 227; Conservative 117; Mismatches 225; Indels 3; Gaps 3;

QY 16 EENFKALVJIAFAQYLQCCPFEDHVKLVNVEVTEFAKTCVADESARCKSLHTLFCKJC 75
Db 40 EMLVGLATIFFAQFVQEAITYKEVQNMVKDVLTVIEKSTGSEQPAGCLENGVSVFLEBIC 99
QY 76 TVATLRITYGEMADCCAKQEPERNECFLOHKKDNP-NLPLRVPEVDVMTAFPHNEEF 134
Db 100 HEEBIEPKYG-LSHCCSQSGSEERHNCFLARKKAAPASIPFPQPEVPTSCKAYEENRETF 158
QY 135 LKLYLVEIARRHPYFVYAPPELLFFAKRYKAAFTCCQAADKAAACLLPKLDELREDEKASA 194
Db 159 MTRVYIARHRPFLAFTILSLAAQYDKIIPCCKAENAVECFQTKAASIKELRESSL 218
QY 195 KQRLKASLQKGERAFKAWARLQRFPAKFAEVSKLVTDLTKVHTCCCHGDLLECA 254
Db 219 LNQHMTVMRQFGARTTRATITVKLSQKFPKANFTETQKLVLDVAHIHEECRGNVLECL 278
QY 255 DORADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDENPADLPSLAADPVEK 313
Db 279 QDAERVVSVYCSQODTLSSKIAESCCKLPTLLELGGCIIHAENDKXPEGLSPNLNRELGR 338
QY 314 DVCKGVAAZADVFLGMLYEVARRHPDYSVLLLRILAKTYETTLKCCAAADPHECYAKV 373
Db 339 DFNQLSREKDLNWARITYSRRHPKLVPTVILRVAKGYQELLEKCSOSRNPLECQDKG 398
QY 374 DDEFKPLVEEPONLIKONCELFEGLEYKFKONALLHRYTKKQVSTPTLVEVSRNLGKV 433

```

```

RESULT 9
Q8UM06 PRELIMINARY; PRT; 624 AA.
AC Q8UM06
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ambystoma texanum (Smallmouth salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8304;
RN [1]
RP SEQUENCE FROM N.A.
RA Haverfield E.V., Urzeil T., Spolsky C.M., Bazartseren B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and
RL Ambystoma texanum.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217182; AAL56645.1; -.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 1.
KW Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 624 SERUM_ALBUMIN.
SQ SEQUENCE 624 AA; 70321 MW; DE08533BF4953EF7 CRC64;

Query Match: 35.0%; Score 1087; DB 13; Length 624;
Best Local Similarity 37.7%; Pred. No. 4.3e-77;
Matches 214; Conservative 102; Mismatches 240; Indels 12; Gaps 5;

QY 14 LGBNFVALVUIAFAQYLQCCPFEDHVKLVNVEVTEFAKTCVADESARCKSLHTLFCK 73
Db 45 IGVEHAKALMALPFSQMLSKCPHHEQVQVRVNVMDIADLCSRGAKHGDGCKSVMTILNE 164
QY 74 LCTVATLRITYGEMADCCAKQEPERNECFLOHKKDNP-NLPLRVPEVDVMTAFPHNEE 132
Db 105 ICKTPEIPEKYPFHGGCKKEDPERHKCFIEHKSTDPKERTYEVKPSPEQICKDHAEIRD 164
QY 133 TFLKKLYIARHRPYPYAPPELLFFAKRYKAAFTCCQ-AADKAAACLLPKLDELREDEKA 191
Db 165 EFLGHYIHKVASSHTTMYPPAILSFTHFDGIVSHCCKDEATVGQCLSEKMPAKKEVEH 224
QY 192 SSARQRLKASLQKGERAFKAWARLQRFPAKFAEVSKLVTDLTKVHTCCCHGDL 251
Db 225 VCQVQKNCVILQNFNERALRASKAHACSKFPHASPVORLTGDGIVHLHQTCGCGDMV 284
QY 252 ECADDRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDENPADLPSLAADPVE 311
Db 285 ACMAERMKLTQTQCEK-----KKCEKXPVLERSECIIVLPNDEKPADLSPEVRYFD 336
QY 312 SKOVCKRYAAKQVFLGMLFVEARRHPDYSVLLLRILAKTYETTLKCCAAADPHECYA 371

```

```

Db 337 DPEVCKRPFKEGGAFMRFLCDYAKIHPEHSAENLRIASGLEKAYKATCCAGFAHNECIA 396
QY 372 KVFDEFKPLVEEPQNLKQNCLEFQEGYKFKQKALLVRYTKKVPQVSTPTLYEVSRLNG 431
Db 397 KEETELRHEIEASKTKKTTCCALEKLGPHFONMIVRYTGILPRSDAFLYTKLT 456
QY 432 KVGSKCKHPEAKMPCAEYLSVVLNOLCVLHEKTPVSRVTCKCTESLVNRRPCFSA 490
Db 457 NIGQCKKLPEDQOMPSEGGLGVFAQICQNKPTPFENEKLAHCCKDSLSFTTPCFAA 515
QY 491 LEVDETVVPKEFNAETTFTHADCTLSEKEROIKKOTALVELVHKHFKATKEQKAWND 550
Db 516 LTVDETVVPATVTAESNFNDEFCTPSEADLQAKQTJXHLVRTHPKITDEQVKTISEK 575
QY 551 FAATVEKCKCAADKCTGFAEEGKXVAA 578
Db 576 FLAMCGGCKADQNECFATBGAJVEA 603

RESULT 10
Q8BK65 PRELIMINARY; PRT; 605 AA.
AC Q8BK65;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Alpha-fetoprotein precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR ENBL; AK076053; BAC36150.1; --
SQ SEQUENCE 605 AA; 67322 MW; 048B7A4A8BC1EA45 CRC64;

Query Match 34.9%; Score 1082; DB 11; Length 605;
Best Local Similarity 35.7%; Pred. No. 1e-76;
Matches 204; Conservative 120; Mismatches 241; Indels 6; Gaps 3;

QY 16 BENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEPAKTCVADESACNCKSLHTLFGDKLC 75
Db 40 EKNVLSATITFTQFVPEATEEE---VNKMTSDVLAAMKKNKSGDGLSGLSVFLDEIC 95
QY 76 TVATLRETYGEMADCCAKQEPNERCEFLQHKDNP-NLPRVVRPEVDVMCTAFHNEET 134
Db 96 HETELSNKYG-LSGCCSGSGVERHQCLLARKKTAPASVPPFPPEPAESCKAHEENRAVF 154
QY 135 LKKYLYEIARRHPFYAPPELLFFAKRYKAAFTCCCAAKACCLPKLDELDEGKASSA 194
Db 155 MNRPIYEVSRRNPWYAPAILSLAAQYDKVVLACCKADNKEECFCCTKRASIAKELREGSM 214
QY 195 KQRLKASLOKGERAFKAWAVARLSQRPFAEFVSKLVTLTKVHTCHCHDLECA 254
Db 215 LNEHVCVIRKPGSNLQATTIILSKLAECCKPMIQGFCIIHAENGVRKEGLSNPQFLGDRN 274
QY 255 DDRAALAKYICENQDSISKKLECECKEPLEKSHCIAEVENDEMPADLPSLAADFESKD 314
Db 275 ODGEKWTYICSOQNTLSKIAECCKPMIQGFCIIHAENGVRKEGLSNPQFLGDRN 334
QY 315 VCKNVAEKDVFLGNFLYEYARRHPDYSVLLLRILAKTYETTLKCCAAADPHECYAKVF 374
Db 335 PAQFSSEKIKFMASFLHEYSRTHNLPVSVILRAKTYQEIILEKCSQSGNLPQCDNLE 394
QY 375 DEFKPLVEEPQNLKQNCLEFQEGYKFKQKALLVRYTKKVPQVSTPTLYEVSRLNGKYG 434
Db 395 KEETELRHEIEASKTKKTTCCALEKLGPHFONMIVRYTGILPRSDAFLYTKLT 454
QY 375 DEFKPLVEEPQNLKQNCLEFQEGYKFKQKALLVRYTKKVPQVSTPTLYEVSRLNGKYG 434

```

```

Db 395 EELQKHIEESQALSQSCALYQTLGDYKQLQNLFLIGYTRKAPQLTSAELIDLTGKMSVIA 454
QY 435 SKCCKHPEAKMPCAEYLSVVLNOLCVLHEKTPVSRVTCKCTESLVNRRPCFSALEVD 494
Db 455 STCCQLSEKSGCGEGMADPIGHLCIRNEASPNVSGISHCCNSYSNRRLCITSLFLAD 514
QY 495 ETVVYKPFNAETTFTHADICTLSEKEROIKKOTALVELVHKHFKATKEQKAWMDFAAF 554
Db 515 ETVASPPSEDFIFHKOLCOAHGKALQTMQOELLINLVKQKPELTEEGLAAVATADFSGL 574
QY 555 VEKCKCAADKCTGFAEEGKXVAAQALGL 585
Db 575 LEKCKCAQDQEVCFTEEGPKLISKTRDALGV 605

RESULT 11
Q8BK56 PRELIMINARY; PRT; 605 AA.
AC Q8BK56;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE Alpha-fetoprotein precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR ENBL; AK076197; BAC36249.1; --
SQ SEQUENCE 605 AA; 67409 MW; CF3509A9EC14611C CRC64;

Query Match 34.7%; Score 1076; DB 11; Length 605;
Best Local Similarity 35.6%; Pred. No. 3.1e-76;
Matches 203; Conservative 120; Mismatches 242; Indels 6; Gaps 3;

QY 16 BENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEPAKTCVADESACNCKSLHTLFGDKLC 75
Db 40 EKNVLSATITFTQFVPEATEEE---VNKMTSDVLAAMKKNKSGDGLSGLSVFLDEIC 95
QY 76 TVATLRETYGEMADCCAKQEPNERCEFLQHKDNP-NLPRVVRPEVDVMCTAFHNEET 134
Db 96 HETELSNKYG-LSGCCSGSGVERHQCLLARKKTAPASVPPFPPEPAESCKAHEENRAVF 154
QY 135 LKKYLYEIARRHPFYAPPELLFFAKRYKAAFTCCCAAKACCLPKLDELDEGKASSA 194
Db 155 MNRPIYEVSRRNPWYAPAILSLAAQYDKVVLACCKADNKEECFCCTKRASIAKELREGSM 214
QY 195 KQRLKASLOKGERAFKAWAVARLSQRPFAEFVSKLVTLTKVHTCHCHDLECA 254
Db 215 LNEHVCVIRKPGSNLQATTIILSKLAECCKPMIQGFCIIHAENGVRKEGLSNPQFLGDRN 274
QY 255 DDRAALAKYICENQDSISKKLECECKEPLEKSHCIAEVENDEMPADLPSLAADFESKD 314
Db 275 ODGEKWTYICSOQNTLSKIAECCKPMIQGFCIIHAENGVRKEGLSNPQFLGDRN 334
QY 315 VCKNVAEKDVFLGNFLYEYARRHPDYSVLLLRILAKTYETTLKCCAAADPHECYAKVF 374
Db 335 PAQFSSEKIKFMASFLHEYSRTHNLPVSVILRAKTYQEIILEKCSQSGNLPQCDNLE 394
QY 375 DEFKPLVEEPQNLKQNCLEFQEGYKFKQKALLVRYTKKVPQVSTPTLYEVSRLNGKYG 434
Db 395 EELQKHIEESQALSQSCALYQTLGDYKQLQNLFLIGYTRKAPQLTSAELIDLTGKMSVIA 454
QY 435 SKCCKHPEAKMPCAEYLSVVLNOLCVLHEKTPVSRVTCKCTESLVNRRPCFSALEVD 494

```

```

Db 455 STCCQLSEKXSCGSGMAD:FIGH:CIRNEASPVNSG:SHCCNSYSNRR:CTISFLAD 514
Qy 495 ETVYKPEKNAETFTFHADICTLSEKEROIKKOTALVELVHKHPKATKEQKAWMDFAAF 554
Db 515 ETVYVPPFSEDKTIFHKDLCOAQKALQTMKQELLNLVKQKPELTTEEQLAAVTADFSGL 574
Qy 555 VEKCKADDKETCFABEGKLVAAQAALGL 585
Db 575 LEKCKAQDOEVCTFEBGPKL:SKTRDALGV 605

RESULT 12
Q8JIA7
ID Q8JIA7 PRELIMINARY; PRT; 400 AA.
AC Q8JIA7
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE A/B over-sized serum albumin (Fragment).
CS Sphenodon puccatus (Bacteria) (Tuatara).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_TaxID=8508;
RN [1]
RP SEQUENCE FROM N.A.
RA Metcalf V.J., Brennan S.C., George P.M., Chambers G.K.;
R "Partial mRNA sequence for tuatara A/B serum albumin."
RJ Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375973; AAV46106.1; -.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
FT NON_TER 1
FT SIGNAL 1
SQ SEQUENCE 400 AA; 45715 MW; 8DE20609657CF753 CRC64;

Query Match 33.7%; Score 1045; DB 13; Length 400;
Best Local Similarity 47.7%; Pred. No. 5.1e-74;
Matches 186; Conservative 80; Mismatches 124; Indels 0; Gaps 0;

Qy 196 QRLKASLQKGFGERAFKAMAVARLSQFFPKAEVSKLVDTTKVHTCCSHGLECAD 255
Db 3 EKHSQGLXSGFERAFCAKLVGLSQKFPKAPFEIHKVLTATKLQKQCHGDMTECLD 62
Qy 256 DRADLAKYICENQDSISSK:KECEKELLEKSHCIAEVDENMPADLPFLAADFVESHKV 315
Db 63 DRVEVMAYICSKQAVFSKIKCCCEKPIVQREVCILQACLEKPAJLPSIAGGYIESTEV 122
Qy 316 CKNYAEAKQVFLGMFLYEVARRHPDYVSWLLRLAKTYETTTLEKCAAADPHECYAKVFD 375
Db 123 CKHYEESKDVFLAHFVYEVSRHPERSSQMLLTGKGYQDTLXCKCTNPEPCYKAGE 182
Qy 376 EFKPLVEEPQNLIKONCELFQGEYKFNALLVRYTKVPQVSTPTLVEVSNLKGVS 435
Db 183 ELARHIOESGELLKTHCSFVTSQCKDPQKMWLVRYTKVPQVPAABELIEISKLTGVG 242
Qy 436 KCGHPEAKMPCADYLSVNLQCVLHEKTPVSDRYVTCCTESLVNRRPCEFALEVDE 495
Db 243 KCCPLSDKRLSCSEKHLSMVLEICRQHEASPVNNHVTTCCTDSYSEMRPCTFLGVDD 302
Qy 496 TYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVHKHPKATKEQKAWMDFAAFV 555
Db 303 SYVPEFCPTFTLDEQLCTAPEARLKKGLTLVKLQKPOIEDEQLKLVDTFHAVE 362
Qy 556 EKCKADDKETCFABEGKLVAAQAALGL 585
Db 363 EKCCQAENKQCEFTSEGEKLTQEGKALLGV 392

RESULT 13

```

```

Q9YGH6
ID Q9YGH6 PRELIMINARY; PRT; 603 AA.
AC Q9YGH6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Rana shqiperica.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Embiibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.
OX NCBI_TaxID=44326;
RN [1]
RP SEQUENCE FROM N.A.
RA Jzzell T., Hotz H.;
RC STRAIN=Busnat; TISSUE=Liver;
RT "Albumin cDNA sequence of Rana shqiperica: evolutionary changes in
RL frog albumins."
DR EMBL; U40452; AAD09358.1; -.
DR HSSP; P02768; 157B.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Signal.
FT NON_TER 1
FT SIGNAL 23
FT CHAIN 24 603 SERUM ALBUMIN.
SQ SEQUENCE 603 AA; 69293 MW; 340D3723FA010C99 CRC64;

Query Match 30.8%; Score 955; DB 13; Length 603;
Best Local Similarity 33.6%; Pred. No. 9.9e-67;
Matches 190; Conservative 119; Mismatches 241; Indels 16; Gaps 3;

Qy 12 KDLGEENFKALVLIAPQYLYQQCPFFDHVKLVNVEVTEFAKTCVADESAENCDKSLHTFG 71
Db 37 KAVGKPAVEKLVLMVAQDFPEKCSLDEHLKVQAKITAEAVDNCCKHPPEAECKKPALEYH 96
Qy 72 DKLTATATLSTTGTGMADCCAKOEPERNBCFLQHKDQNPMLPLVR-----REVDMCT 125
Db 97 DVCKEEDIDQLYPTWTTCCGKAEEATKCFYEHRE-----VRVEYKFNIEESCK 148
Qy 126 AFHNEETFLKYLVEIARHPYFYAPELFFPAKRYKAAFTCCQADKAAACLPKLDEL 185
Db 149 EKHSHQGRAFSYLSN:AKGRSKLYPPAVLGPAIQYNEITTECCAAEDKAKCKGERMPQV 208
Qy 186 RQEGKASSAKQRLKASLQKGFGERAFKAMAVARLSQFFPKAEVSKLVDTTKVHTEC 245
Db 209 KKLNYLEDKHKQCRVLKSFPERVSQALTLVQVSQFGNAKYDDVEKVTFIEAHLNEDC 268
Qy 246 CHGDLLECADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVDENMPADLP 305
Db 269 CKGDAVCEMTIERMEATHICLAKELKSSLSJSCCAKGVLETPCICALPNEE--PDLP 326
Qy 306 AADPEVSKQVCKNYAEAKQVFLGMFLYEVARRHPDYVSWLLRLAKTYETTTLEKCAAAD 365
Db 327 LKEYEEDHVCENYQKDKRYLAHFTHDYSRSHQESSPQSLVSRGFEMLLLEKCCASAN 386
Qy 366 PHECYAKVDFEKPFLVEEPQNLIKONCELFQGEYKFNALLVRYTKVPQVSTPTLVE 425
Db 387 SAECKDAPKLLEAAKNEEIEISKQNCGALEKGLGNDFYQLLVRYFGKMPQVTAOTLVE 446
Qy 426 VSNLKGVSCKCHPEAKMPCADYLSVNLQCVLHEKTPVSDRYVTCCTESLVNRR 485
Db 447 LTGRMAKIGVYCCGLPDNKKQPCAEKLDILLGEMCEREKKTFINDNVHHCVDVSNRR 506
Qy 486 PCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVHKHPKATKEQK 545
Db 507 PCFTKLPGLYKAYEAPVWDESKLHFTADMCKGSADDQLKT:KLVLVLFELQKPTCGKEKLT 566

```



QY 532 LVKHKFKATKEQLKAVMDFAAFVEKCKCKXADDKETCFEAEKGKLYAASQAA 583  
DB 354 VVKCRPAITHEQLKAVITDFYGVWEKCKCHGENHEACFLAEGFCLVQRTCAA 455

Search completed: October 27, 2003, 15:21:05  
Job time : 108 secs